DOCUMENT RESUME

ED 054 826 LI 003 102

AUTHOR Chang, Yean--Hsi

TITLE The Implementation of a Relational Document Retrieval System.

SPONS AGENCY Joint Services Electronics Program, Fort Monmouth, N.J.; Office of Education (DHEW), Washington, D.C.

PUB DATE Sep 70

CONTRACT OEC-1-7-071213-4557

NOTE 58p.; (9 References); Thesis submitted in partial fulfillment of requirements for degree of Master of Science in Electrical Engineering, Graduate College of University of Illinois, Urbana-Champaign

EDRS PRICE MF-$0.65 HC-$3.29

DESCRIPTORS Data Bases; *Information Retrieval; *Information Systems; Programing Languages; *Relevance (Information Retrieval); *Search Strategies

ABSTRACT

A relational document retrieval system is developed that is primarily concerned with hierarchic tree structures and lateral links. The lateral linkages are relationships that cannot be represented within the tree structure representation. A number of annotated examples are given along with a detailed description of the relational data structure and the programs that manipulate and search this structure. The system was programmed in a string manipulating language designed for information retrieval purposes. (Author)

ED054826

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE IMPLEMENTATION OF A RELATIONAL DOCUMENT RETRIEVAL SYSTEM

BY

YEAN-HSI CHANG

This work was supported by the Office of Education under Contract OE C-1-7-071213-4557; auxiliary support was provided by Joint Services Electronics Program (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DAAB-07-67-C-0199.

Reproduction in whole or in part is permitted for any purpose of the United States Government.

This document has been approved for public release and sale; its distribution is unlimited.

LI 003 102

THE IMPLEMENTATION OF A RELATIONAL DOCUMENT RETRIEVAL SYSTEM

BY

YEAN-HSI CHANG
B.S., Chinese Naval College of Technology, 1953

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical Engineering,
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1970

Urbana, Illinois

LI 003 102

THE IMPLEMENTATION OF A RELATIONAL DOCUMENT RETRIEVAL SYSTEM

Yean-Hsi Chang, M.S.
Department of Electrical Engineering
University of Illinois, 1970

A relational document retrieval system is developed that is primarily concerned with hierarchic tree structures and lateral links. The lateral linkages are relationships that cannot be represented within the tree structure representation. A number of annotated examples are given along with a detailed description of the relational data structure and the programs that manipulate and search this structure. The system was programmed in ISL, a string manipulating language designed for information retrieval purposes.

ACKNOWLEDGMENT

The author wishes to express his gratitude to Professor R. T. Chien for his guidance and criticism in the preparation of this thesis, and especially to Mr. Fred Stahl, Ph.D. candidate of the Department of Computer Science, for teaching the operation of the CDC 1604 computer, the ILLAR and ISL systems, and his valuable suggestions. Thanks also to Mr. Nickolas Esser for the compiled material in his master's thesis, and to Mr. Karl Kelley for his help in debugging the program. Thanks are extended to Mrs. Sandy Bowles and Mrs. Sherry Kallembach for the typing of this manuscript.

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
II.	THE USE OF RELATIONAL INFORMATION IN DOCUMENT RETRIEVAL	3
III.	THE RELATIONAL STRUCTURE DESIGN	9
IV.	MAJOR FUNCTIONS NECESSARY FOR THE USE OF RELATIONAL STRUCTURES	14
V.	CONCLUSION	15
REFERENCES		16
APPENDIX		
A.	DESCRIPTIONS OF PROGRAMS	17
B.	USER'S MANUAL	22
C.	SAMPLE OF THE HIERARCHIC TREE STRUCTURE USED TO TEST THE PROGRAM	25

I. INTRODUCTION

Document retrieval systems have been intensively studied in recent years. One of the more interesting aspects in this area is that of file structuring. One particular file structure may be found to be very effective in a particular system unacceptable in another system. A review of some of the well known document retrieval systems shows that there are many problematic aspects still remaining with respect to the file structure. The AESOP system [7] uses hierarchic tree structure to handle the category titles and tables to handle the items in the categories. The BOLD system [8] uses list subject category and matraces to show the relation between documents and index terms. The SMART system [9] uses an inverted file of clusters. But no one has considered relations between documents, such as: two documents could have the same descriptors and yet be discussing entirely different subjects, or conversely if two documents are concerned with the same topic but have entirely different descriptors; also the relationship between certain documents could be conditional upon the intent of the user. These relationships can not be expressed clearly by a simple list or a simple hierarchic tree structure. In view of this fact, Esser [1] has done some theoretical work on hierarchic tree structuring together with lateral links (LL) relating the terms used in the field of coding theory. The work reported here represents the development and implementation of the computational aspects of a relational data file for document retrieval employing the tree structuring a lateral-link concepts.

In order to give an overall picture of a system that employs these techniques described previously and shows the effectiveness of such an approach

a number of annotated examples will be given in Chapter II. Following these examples a detailed description of the relational data structure developed is given in Chapter III. In Chapter IV a formal description of all the features available in system is given.

The Appendix contains a description of the programs, a user's manual and a sample tree structure from the relational model used.

This document retrieval system was programmed in ISL, a string manipulating language designed for information retrieval [2,3,4,5,6].

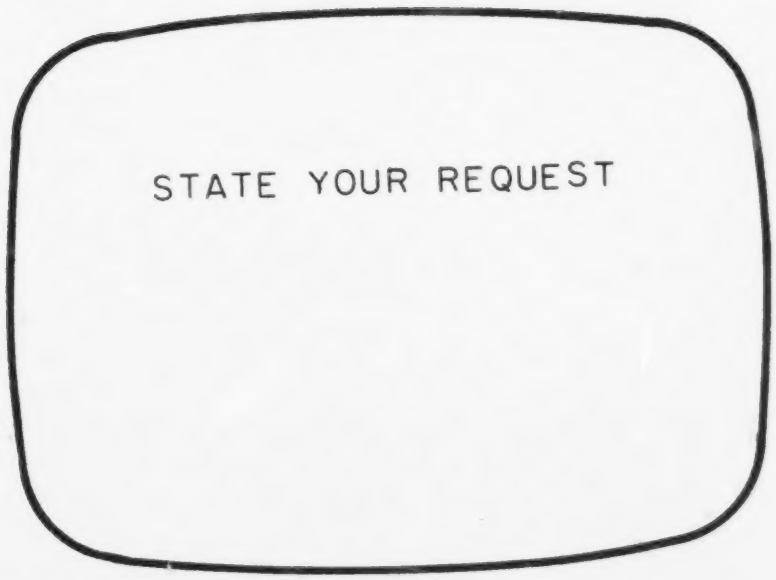
II. THE USE OF RELATIONAL INFORMATION IN DOCUMENT RETRIEVAL

The relational document retrieval system implemented in this paper is one which uses all the available relations between items of the data base to give the user as much information as possible. It is also designed in such a way that the user needs no technical background in order to do the retrieval. The examples given in this chapter show the system's actual performance. The detail step-by-step description of how to operate the system is given in Appendix B.

The data base area used for this implementation is from the area of coding theory. The names are either the categories or the terms generally used in coding theory texts, e.g., cyclic codes, optimum codes, decoding algorithms, binary codes, etc. Each node in the tree structure represents a topic area.

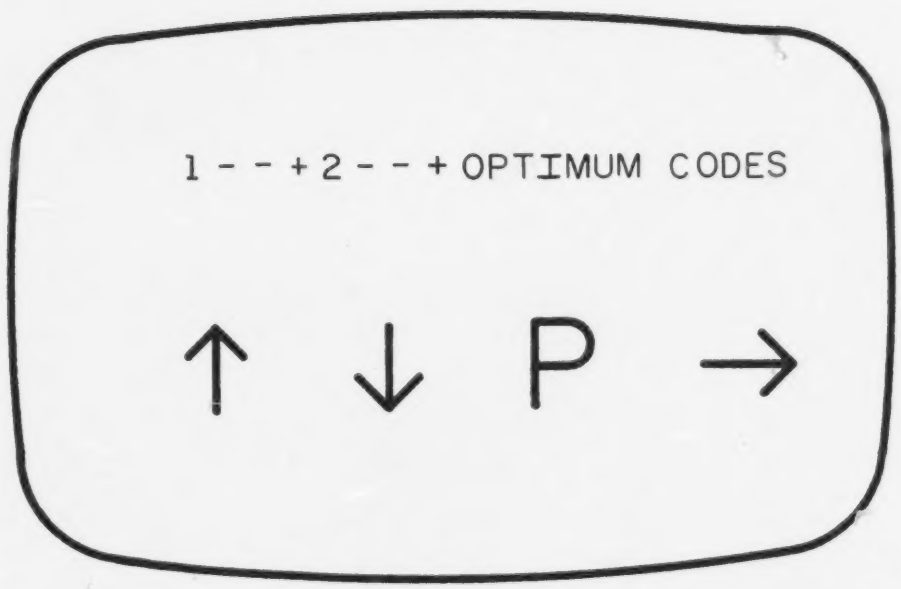
If the user has in mind a topic in coding theory and he wishes to know more about it, then this program can be used. Suppose that a user has "OPTIMUM CODES" in mind. After the computer has been set up by following the procedure given in Appendix B, the statement "STATE YOUR REQUEST" will appear on the display console as in Figure 1 and will also be typed on the console typewriter. Now, he can type "OPTIMUM CODES" followed by a carriage return. At this point in the program "ABORT" may be typed instead and the program control will be passed to the monitor.

The topic that has been typed will be shown in the display console. The program can be restarted again by typing "ERASE" or the program will continue by typing "GO". After typing "GO", the display console will appear as in Figure 2. Pressing the right-most arrow on the screen with the lightpen,



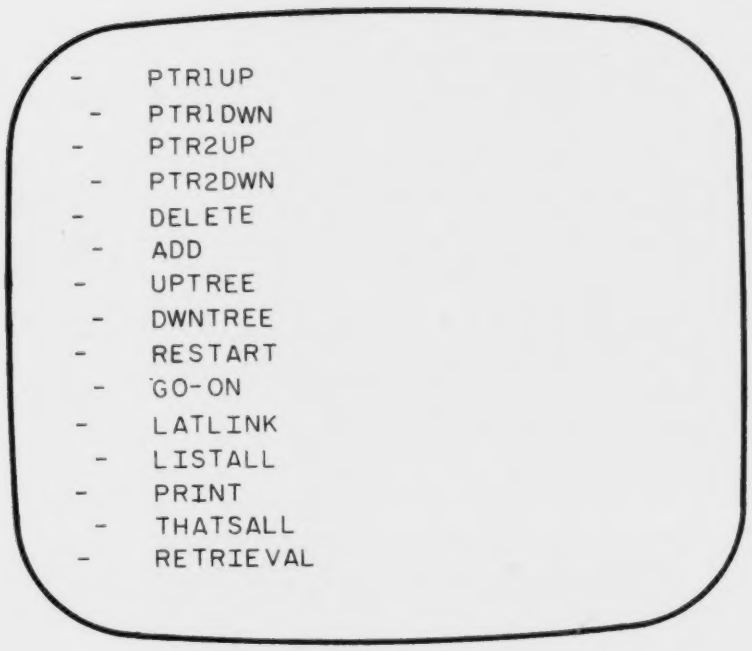
FP-2369

Figure 1.



FP-2370

Figure 2.



FP-2371

Figure 3.

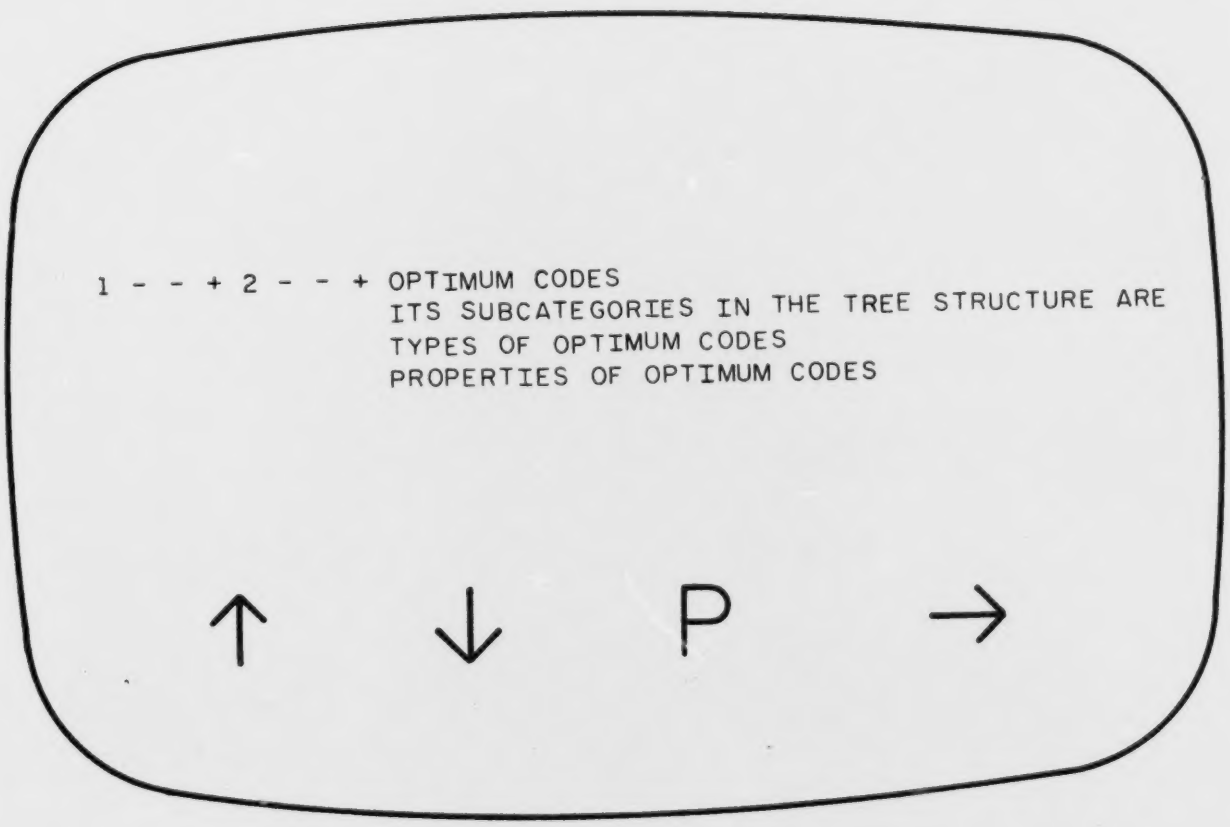


Figure 4.

FP-2372

the display console will appear as shown in Figure 3. There are 15 commands, the functions of which are explained in detail in Appendix B. If "DWNTREE" is selected with the light-pen, the display console will be as shown in Figure 4.

Suppose that the user wants to know the optimum codes under high rate transmission; he can press the right-most arrow with the light-pen to get Figure 3 shown on the display console again and select "ADD" with the light-pen. By typing "HIGH RATE" followed by a carriage return Figure 5 will be shown if "HIGH RATE" is available in the data file (in this case, it is).

By pressing the right-most arrow and selecting the "PTR2DWN" with the light-pen four times until the picture as shown in Figure 6 is shown on the display-console screen.

Pressing the right-most arrow and selecting "LATLINK" with the light-pen will then present the user the information shown in Figure 7.

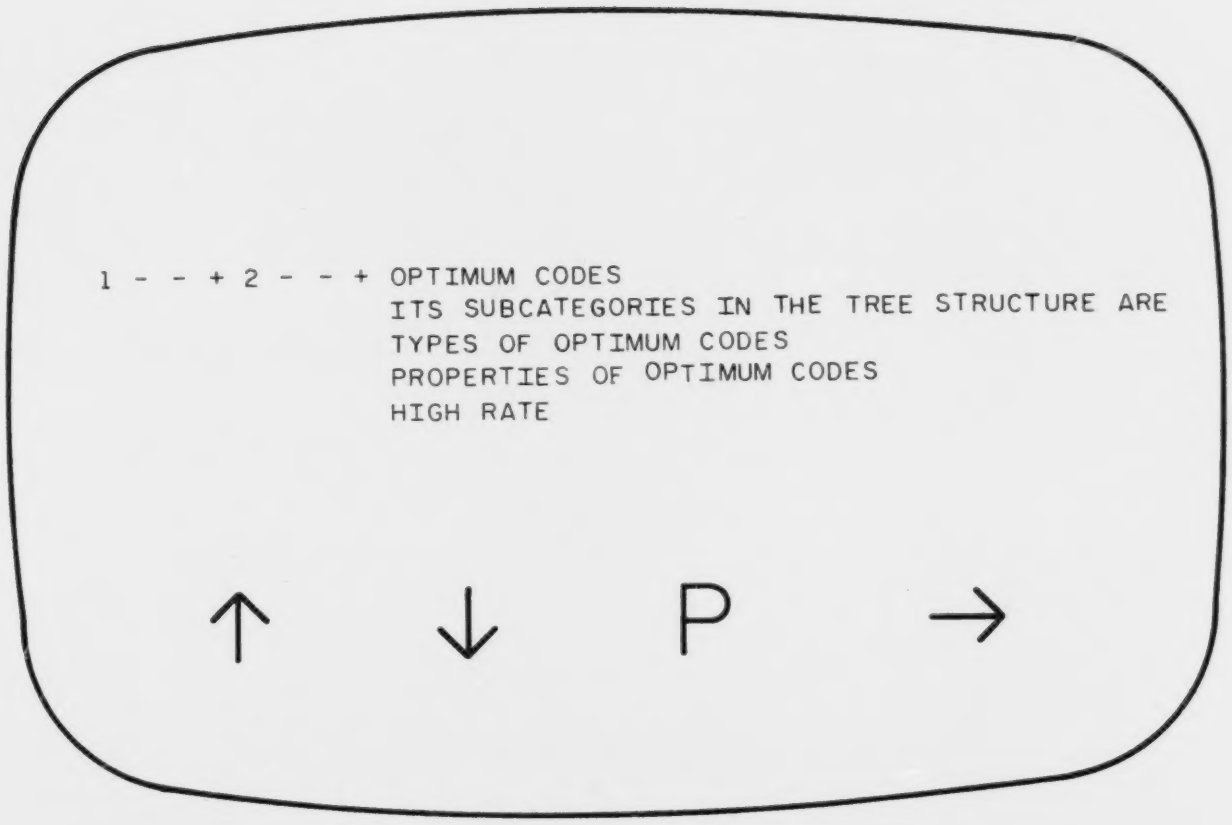
If he presses the right-most arrow with the light-pen and selects "GO-ON", he will see the picture shown as in Figure 8. That means there is only one node whose name is "OPTIMUM CODES".

The user may start the retrieval process* if he feels that the topic names found are enough for his interests simply by pressing the right-most arrow and selecting "RETRIEVE" with the light-pen.

If he selects "LISTALL", he will have the whole list shown again. By selecting "PRINT", he will have the whole list printed. But, if he selects "RESTART", he will erase the whole list and start the program all over again.

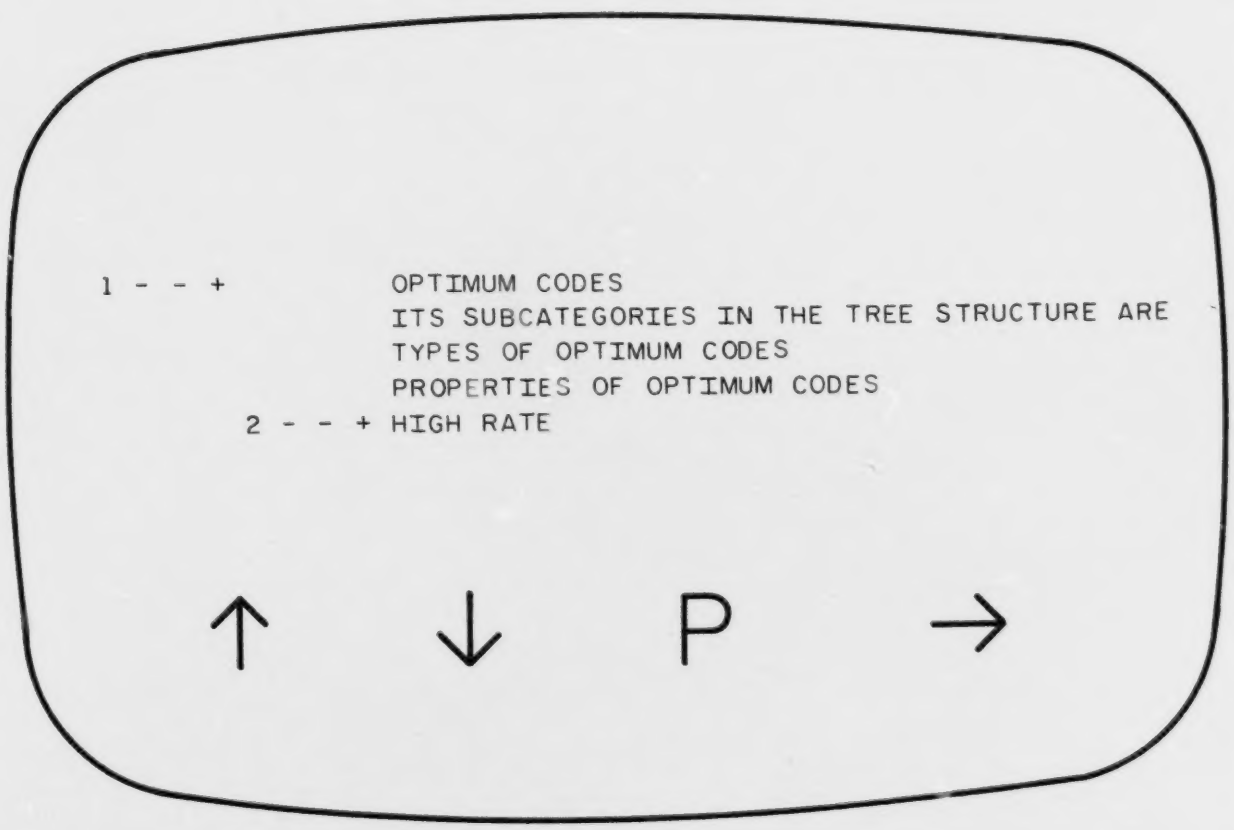
---

*Program for retrieval process is not included in this implementation.



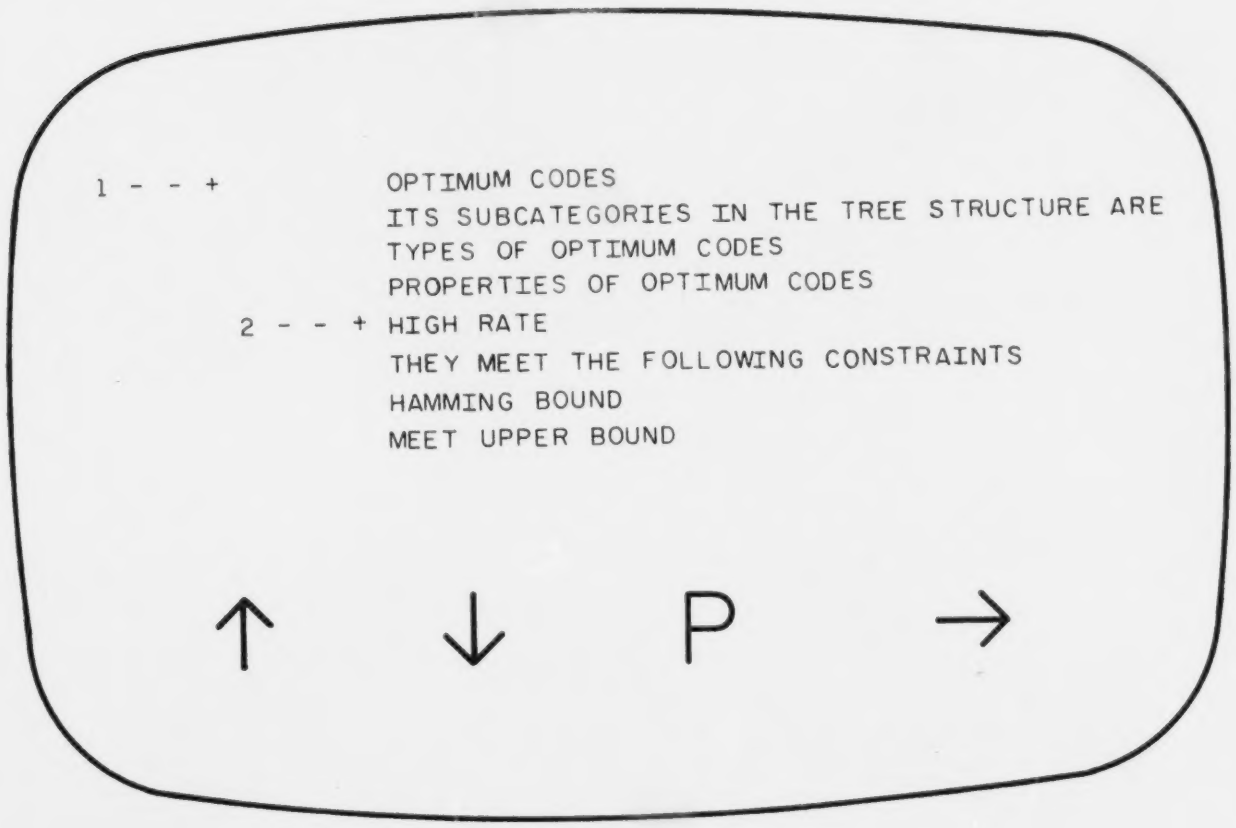
FP-2375

Figure 5.



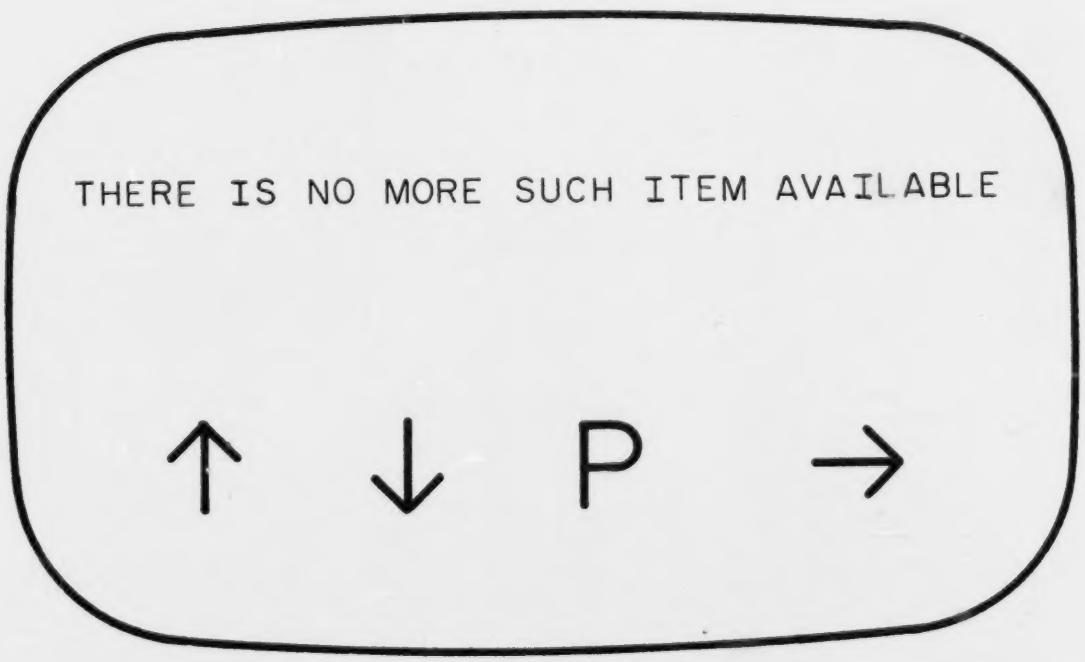
FP-2374

Figure 6.


1 - - +
OPTIMUM CODES
ITS SUBCATEGORIES IN THE TREE STRUCTURE ARE
TYPES OF OPTIMUM CODES
PROPERTIES OF OPTIMUM CODES
2 - - + HIGH RATE
THEY MEET THE FOLLOWING CONSTRAINTS
HAMMING BOUND
MEET UPPER BOUND
↑ ↓ P →

FP-2376

Figure 7.



FP-2373

Figure 8.

III. THE RELATIONAL STRUCTURE DESIGN

In order to best illustrate how the relational information is stored, a representation of the subtree given in Figure 10 will be given. Three structure representations are described here.

A. Node Representation

Each model node consists of two parts as shown in Figure 9. In part one, there are seven fields. Field 1, "NEXT", contains the word address of the next node whose name starts with the same character as this node; field 2, "END", contains a pointer pointing to the byte address of the end of the character string of this node; field 3, "TYPE", contains the information which indicates the type of lateral link function that applies to this node; field 4, "I#", contains the item number of this node; field 5, "UPPER" contains the item number of the node which is at one level up in the hierarchic tree structure; field 6 and 7, "LOWER 1" and "LOWER 2", contain the item numbers of the first and last nodes, respectively, which are at one level down in the hierarchic tree structure. Part two of the node is a variable length string that contains the name of the topic represented by this node.

Figure 11 shows how this information would look for the node given in Figure 10.

B. Lateral Link Representation

These structures consist of three to eight fields depending on which type of lateral link function applies. Field 1 contains the item

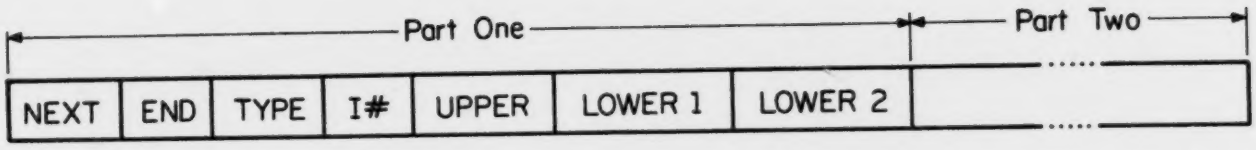


Figure 9.

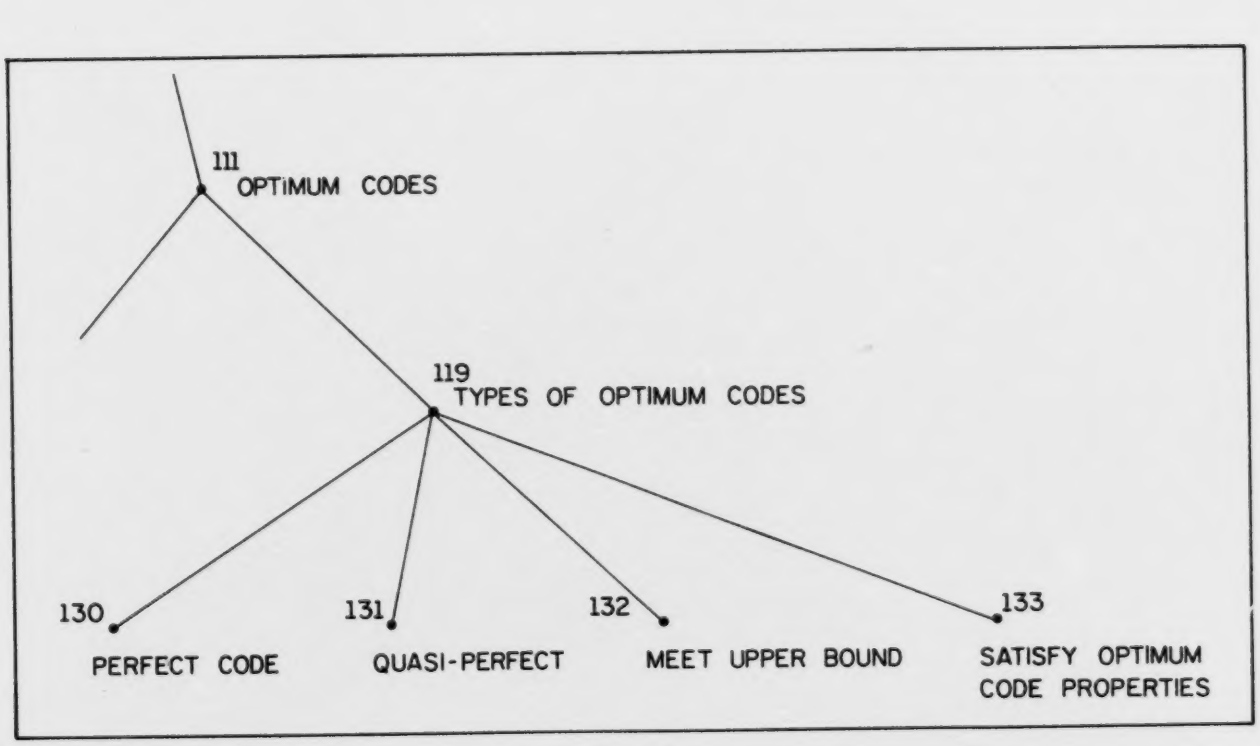


Figure 10. Part of the tree structure about the node whose name is "TYPES OF OPTIMUM CODES".

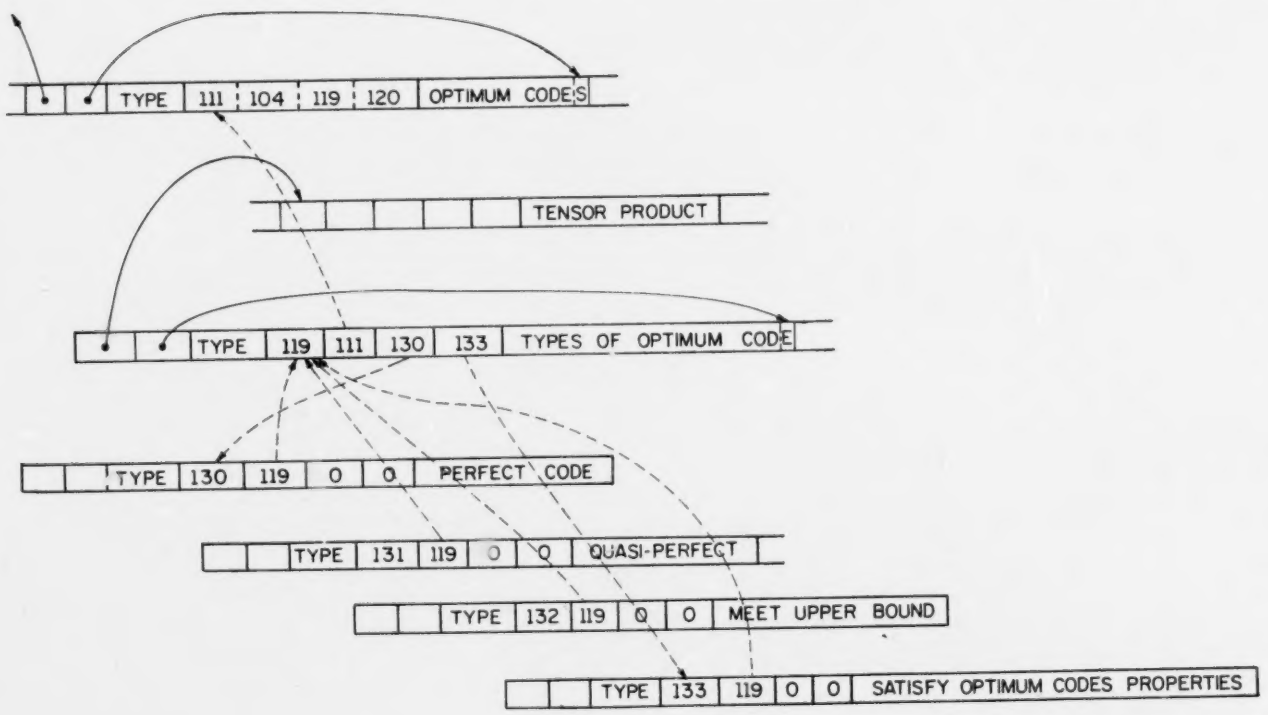

TYPE 111 104 119 120 OPTIMUM CODES
TENSOR PRODUCT
TYPE 119 111 130 133 TYPES OF OPTIMUM CODE
TYPE 130 119 0 0 PERFECT CODE
TYPE 131 119 0 0 QUASI-PERFECT
TYPE 132 119 0 0 MEET UPPER BOUND
TYPE 133 119 0 0 SATISFY OPTIMUM CODES PROPERTIES
FP-2358

Figure 11. Relations between nodes.
——————> related by pointer
– – – –> related by item number

number of the major node, and the rest of the fields contain item numbers of those nodes which required by the lateral link function. For example, for the node "OPTIMUM CODES", the lateral link representation will be as shown in Figure 12, since the lateral link function LLB9 requires 7 item numbers.

C. Representation of Synonyms

These structures have the same form as the regular node representation. The differences are:

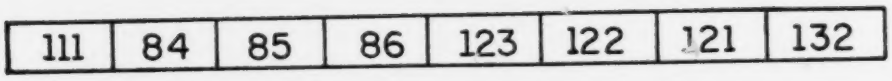
1) Field 3 in part one indicates the synonym property.

2) Field 5 in part one contains the item number of its synonym node.

3) Field 6 and 7 in part one contain zeros, and

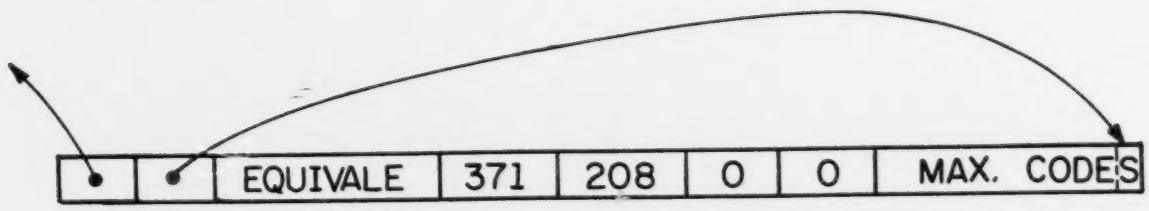
4) Part two contains the name of the item which is not in the tree structure but is a synonym.

For example, "MAX. CODES" is not in the tree structure, but its synonym "MAX. DISTANCE SEPARABLE CODES" is. So its representation is as shown in Figure 13.



FP-2359

Figure 12. Lateral link representation for node "OPTIMUM CODES".



FP-2357

Figure 13. Synonym representation for "MAX. CODES".

IV. MAJOR FUNCTIONS NECESSARY FOR THE USE OF RELATIONAL STRUCTURES

The retrieval action of this program is initiated whenever the name of a topic or the leading portion of it is given. There are four major functions in this implementation.

A. The node is located by its name or the leading portion which is given. Nodes which have the same name or the leading portion, or which are related to the node found either by the hierarchy tree structure or by the lateral link functions will also be suggested, printed, and shown on the display console.

B. In any displayed list of nodes two movable pointers are provided which can move up and down independently for use in performing manipulative actions.

C. A lateral link function will be performed upon command on the nodes pointed to by the pointer 1 (PTR 1) if one record is required by that lateral link function, or on the two records pointed to by the pointers (PTR 1 and PTR 2) if two records are required by a lateral function. Lateral functions which require more than two records will be performed on records in the whole list.

D. Nodes which are not in the tree structure but have synonyms will be located. Its synonym will be printed and displayed as described in A.

V. CONCLUSION

The work presented in this paper is concerned with the hierarchy tree structure and lateral link functions in relational document retrieval systems. Fast access and processing have been carefully considered. The flexibility of this implementation is such that it is possible to apply these principles equally well to any other relational document retrieval data base which possesses hierarchy tree structures and lateral link function relations without any modification of the program.

REFERENCES

1. Esser, N. M., Jr., A Relational Structure for Document Retrieval, Coordinated Science Laboratory Report R-469, University of Illinois, 1970.

2. Kelley, K. C., Ray, S. R., and Stahl, F. A., INFORMATION SEARCH LANGUAGE File No. 735, Dept. of Computer Science, University of Illinois, Urbana, Illinois, September, 1967.

3. Kelley, K. C., Ray, S. R., Stahl, F. A., ISL - A String Manipulating Language, Report R-407, Coordinated Science Laboratory, University of Illinois, Urbana, Illinois, February, 1969.

4. Carroll, D. E., Chien, R. T., Kelley, K. C., Preparata, F. P., Reynolds, P., Ray, S. R., Stahl, F. A., An Interactive Document Retrieval System, Report R-309, Coordinated Science Laboratory, University of Illinois, Urbana, Illinois, December, 1968.

5. Chien, R. T., Ray, S. R., Stahl, F. A., ISL - A New Programming Language for Information Retrieval, Report R-449, Coordinated Science Laboratory, University of Illinois, Urbana, Illinois, December 1969.

6. Chien, R. T., Ray, S. R., Stahl, F. A., "ISL - A New Information Retrieval Language," in Proceedings of the Third International Symposium of Computer and Information Science, edited by Julius Tou, Academic Press, New York, (to appear) 1970.

7. Bennet, E., Haines, E., Summers, J. AESOP: A prototype for outline user control of organizational data storate, retrieval and processing. AFIPS Conf. Proc. Vol. 27, part 1. 1965.

8. Burnaugh, H. P., Introduction to the BOLD Retrieval System, SDC, Santa Monica, Calif. Oct. 1965 (TM-2306/000/03).

9. Ide, E., Salton, G., Interatcive Search Strategies and Dynamic File Organization in Information Retrieval, Report ISR-16, Cornell Univ., 1969.

APPENDIX A

DESCRIPTIONS OF PROGRAMS

Appendix A contains the following information in the order listed below:

(1) MAIN PROGRAM

(2) SUBPROGRAMS

(3) FLOW CHARTS

1. Main Program

This implementation has been programmed on the Control Data 1604 computer at the Coordinated Science Laboratory of the University of Illinois. The main program and the subprograms comprising this implementation were written in the Information Search Language (ISL) [2,3,4,5,6].

There are three portions of the main program in the implementation which will be briefly described here.

a. The first portion of the main program serves to set up the files in the computer memory in the appropriate sequential form. For easy and fast access, directories are built in which the leading addresses of groups and the addresses with respect to item numbers are kept. Twenty-two lateral function directories are also built,each for one of the twenty-two categories of lateral link functions.

b. The second portion of the main program accepts commands from the user and makes a list of the required topic and synonyms and then suggests topics which may be of further interest to the user.

c. The remaining portion of the main program performs further manipulations of the information and executes the desired lateral link functions upon the user's demand.

2. SUBROUTINES

There are 13 subprograms:

a. SETUP: This subprogram arranges the records in the appropriate sequence and sets up the directories and the links.

b. SEARCHA: This subprogram analyzes a given character string and locates where in the relational file it is located.

c. SEARCHB: This subprogram locates a given character string in the portion of the relational file after a given address.

d. COMPARE: This subprogram compares two character strings of a given length to see if they are the same.

e. LFTN: This subprogram serves to perform the 22 lateral functions, to suggest new related nodes or to suggest the deleted of the node found.

f. SETLL: This subprogram puts all the item numbers which will be used in the particular lateral function into the corresponding directory.

g. MOVPTR: This subprogram moves the pointer specified from position 1 (pos1) to position 2 (pos2).

h. CHKINO: This subprogram checks for the item number specified to see if it is in the list DIRECT. The value 0 in accumulator A indicates success. Fail otherwise.

i. GETINO: This subprogram gets the item number of the topic in the record with leading address given and stores it at ITNO.

j. MOVTV1: This subprogram puts the statement with given leading address and ending address in the scope buffer and adjusts the list DIRECT.

k. MOVTV2: When the leading address of the record is given, this subprogram moves the topic title of that record to the scope buffer and sets both lists DIRECT and DIRTTV.

l. GETRT1: This subprogram moves the statement between the two given addresses to the scope buffer. Then, for the record with leading address at RCDADR, it finds all the related records and puts their topic titles in the scope buffer.

m. GETRT2: This subprogram gets the item numbers pointed to by both pointer 1 (1--*) and pointer 2 (2--*), and checks to see if the condition of the lateral link function involved is satisfied. On success, it moves the statement between the two given addresses and the related topic names to the scope buffer.

3. FLOW CHARTS

In the flow charts, the following symbols are used:

a. b : : Index designator,

b. B^b : : Deisgnated index register,

c. ( ) : Contents of a register or storage location,

d. (a1,a2) : Contents between byte locations a1 and a2.

APPENDIX B

USER'S MANUAL

1. To start the program:

   (a) Mount tape C-140 (ISL MASTER) on logic unit 1, and tape F-137 (MAIN) on logic unit 2.

   (b) Mount tape F-138 (DATAMAIN) on logic unit 8.

   (c) Press "AUTO-LOAD" button to initiate the program.

   (d) Type "LOAD,2" followed by a carriage return.

   (e) Type "GO,MAIN" followed by a carriage return.

2. To start the retrieval process:

The program will be ready when "STATE YOUR REQUEST" is typed on the typewriter, and at the same time shown on display console.

Type a topic name or its leading portion which you are interested in. Follow this by a carriage return. What you have typed will also be shown on the display console.

Another "carriage return" will give you the opportunity to terminate the program by typing "ABORT" or change the name by typing "ERASE"; and then a "carriage return".

Your command will be accepted each time the display console is showing information. There are 15 simple commands available:

   (a) PTR1UP: To move pointed 1 (1--*) up the list one step.

   (b) PTR1DWN: To move the pointer 1 down the list one step.

   (c) PTR2UP: To move pointer 2 (2--*) up the list one step.

   (d) PTR2DWN: To move pointer 2 down the list one step.

   (e) DELETE: To delete the topic title pointed to by pointer 1.

(f) ADD: To add a new title at the end of the list.

(g) UPTREE: To find the title of the node one level higher than the one pointed to by pointer 1 in the tree structure.

(h) DWNTREE: To find the title of the node one level lower than the one pointed to by pointer 1 in the tree structure.

(i) RESTART: To erase the whole list and start the program over again.

(j) GO-ON: To keep the list in the memory and go on the retrieval process.

(k) LATLINK: To apply the lateral function on the topics pointed to by pointer 1 and also pointer 2 or of the whole list if function requires.

(l) LISTALL: To restore the part of the list kept in the memory and show them all on the TV screen.

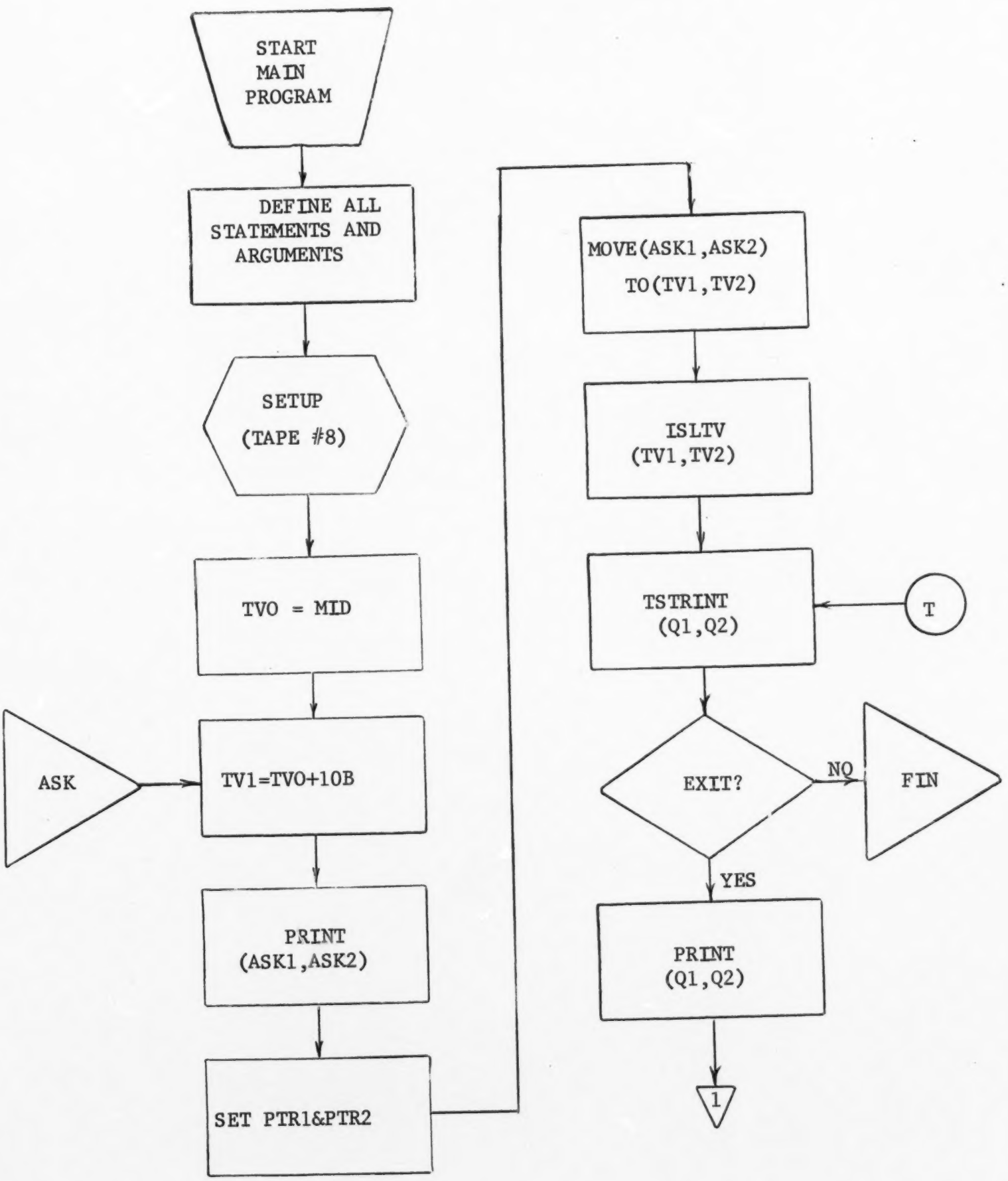
(m) PRINT: To print out whatever shown in the TV screen.

(n) THATSALL: To terminate the program.

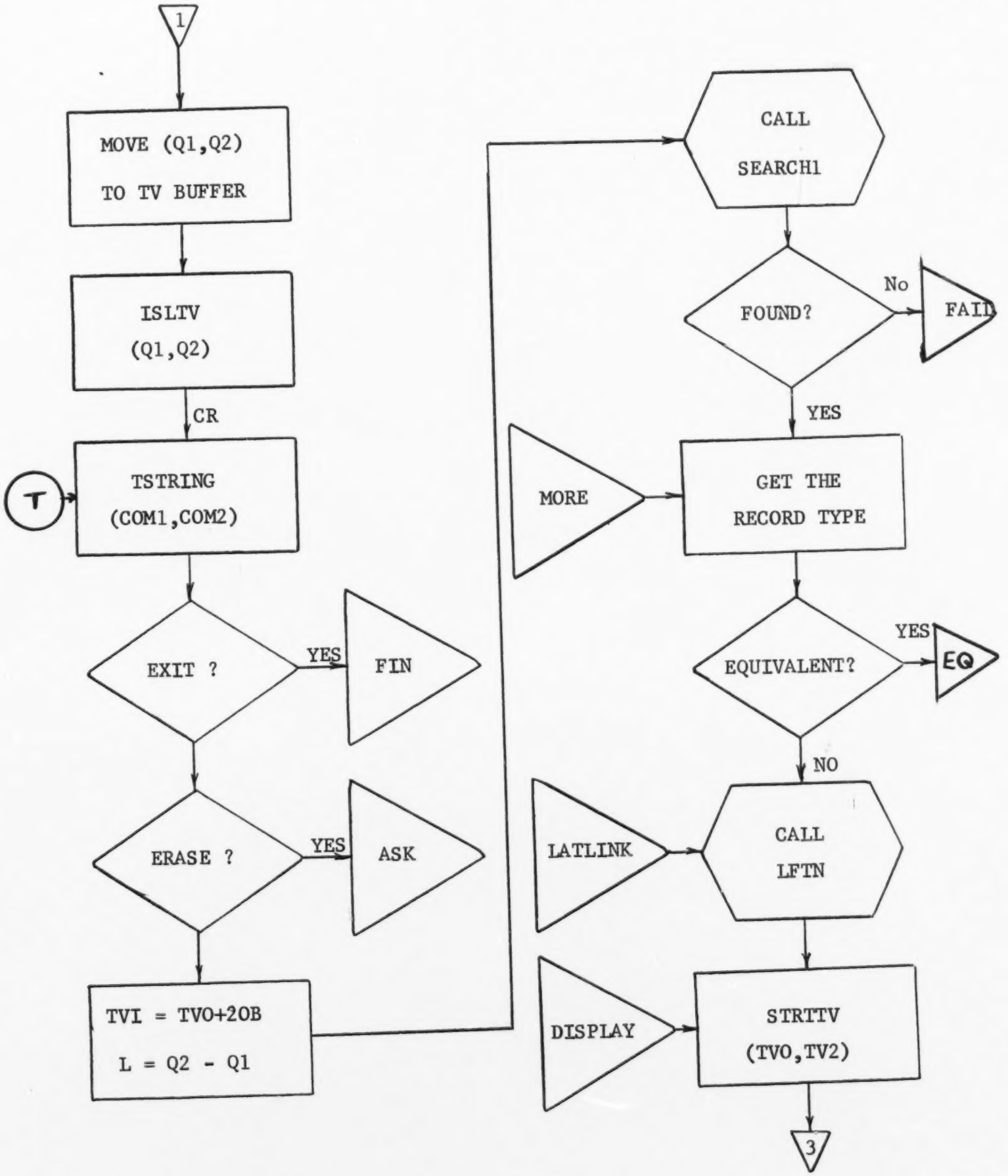
(o) RETRIEVAL: (Not included in this program).

APPENDIX C

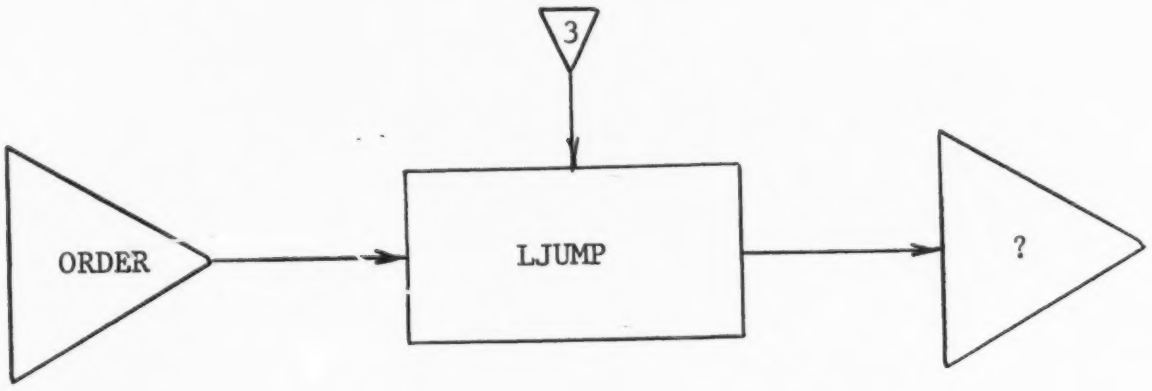
SAMPLE OF THE HIERARCHIC TREE STRUCTURE USED TO TEST THE PROGRAM


T : Typewriter Input
START MAIN PROGRAM
DEFINE ALL STATEMENTS AND ARGUMENTS
SETUP (TAPE #8)
TVO = MID
ASK
TV1=TVO+10B
PRINT (ASK1,ASK2)
SET PTR1&PTR2
MOVE(ASK1,ASK2) TO(TV1,TV2)
ISLTV (TV1,TV2)
TSTRINT (Q1,Q2)
T
EXIT?
NO
FIN
YES
PRINT (Q1,Q2)
1

MAIN PROGRAM CONTINUE



MAIN PROGRAM CONTINUE



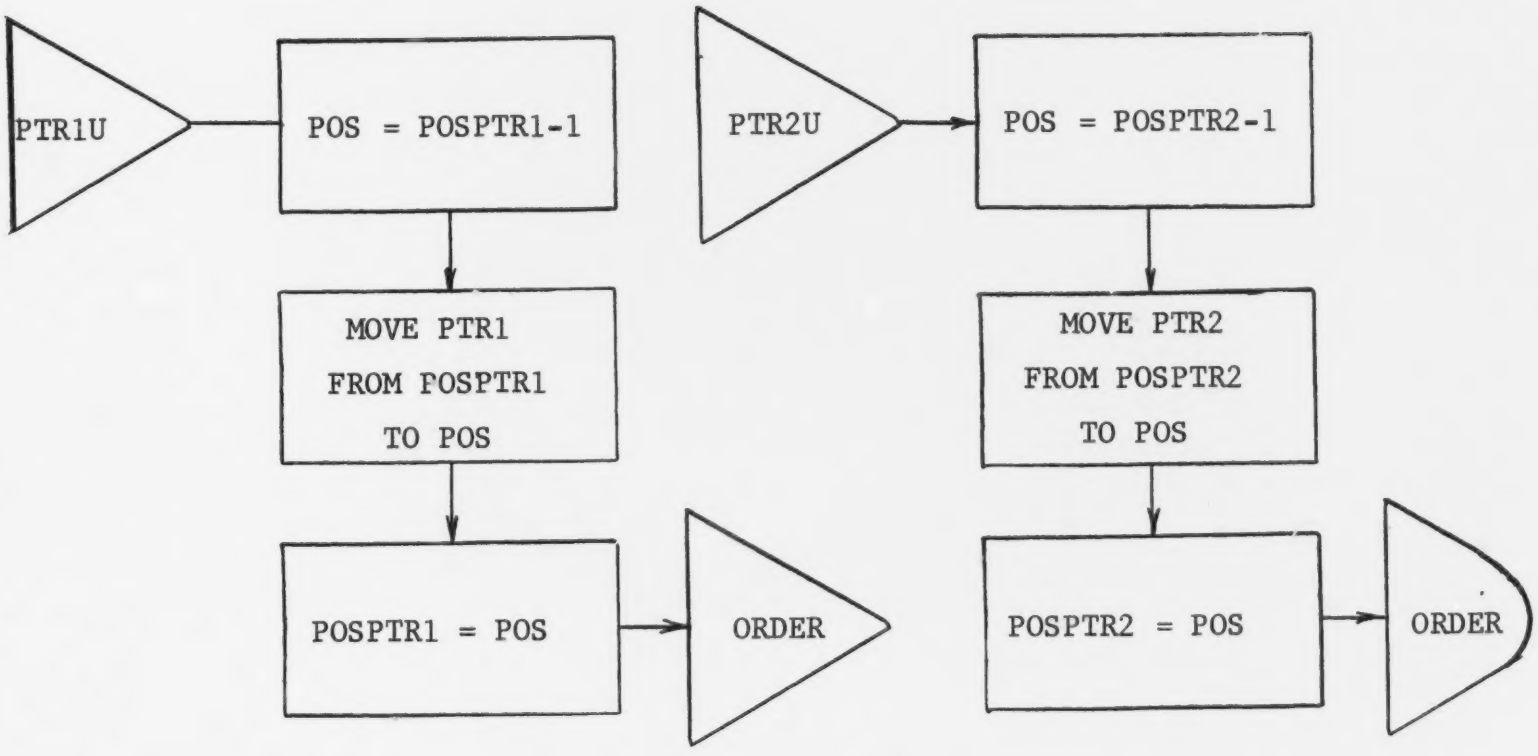
SELECT WITH LIGHT-PEN

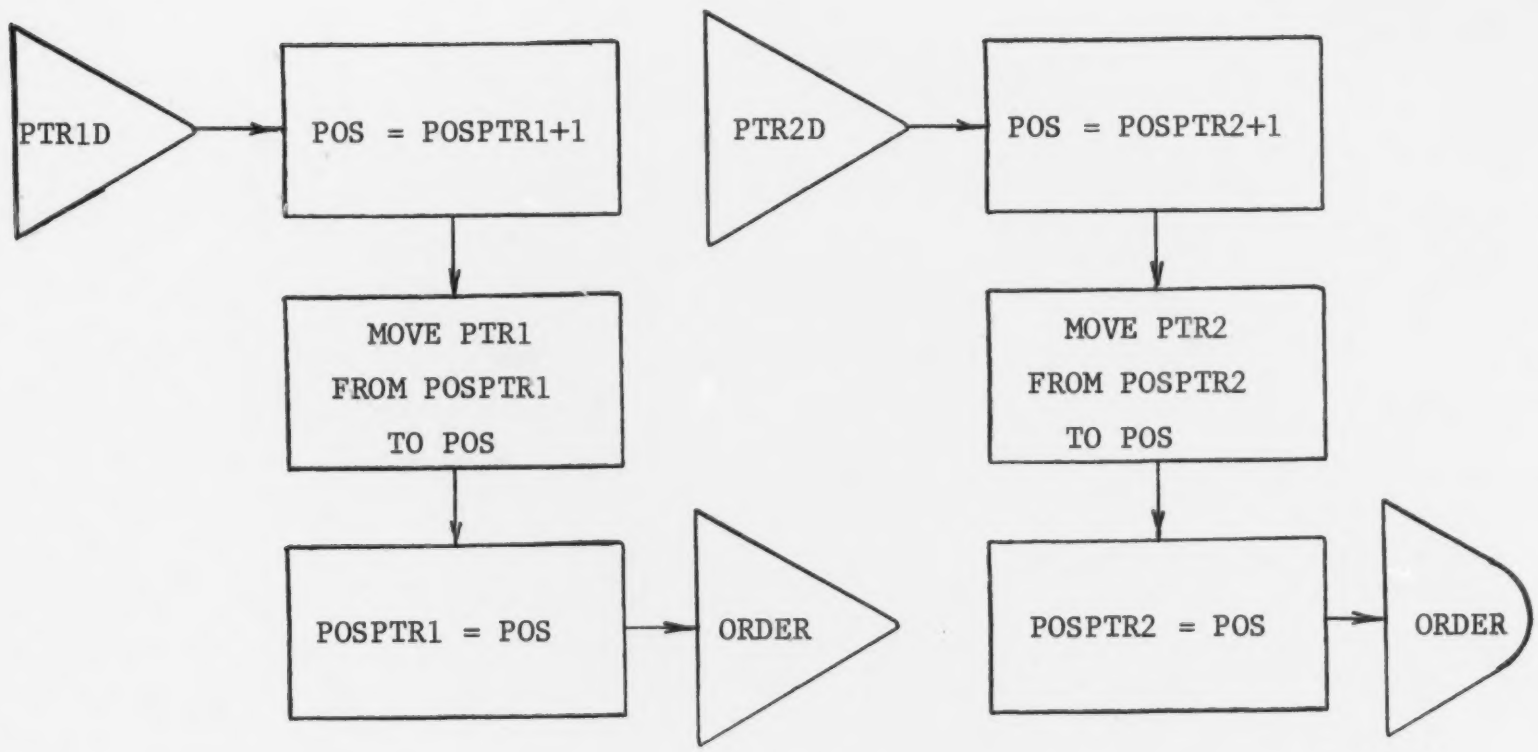
1. - PTR1UP
2. - PTR1DWN
3. - PTR2UP
4. - PTR2DWN
5. - DELETE
6. - ADD
7. - UPTREE
8. - DWNTREE
9. - RESTART
10. - GO-ON
11. - LATLINK
12. - LISTALL
13. - PRINT
14. - THATSALL
15. - RETRIEVE

GOTO ?

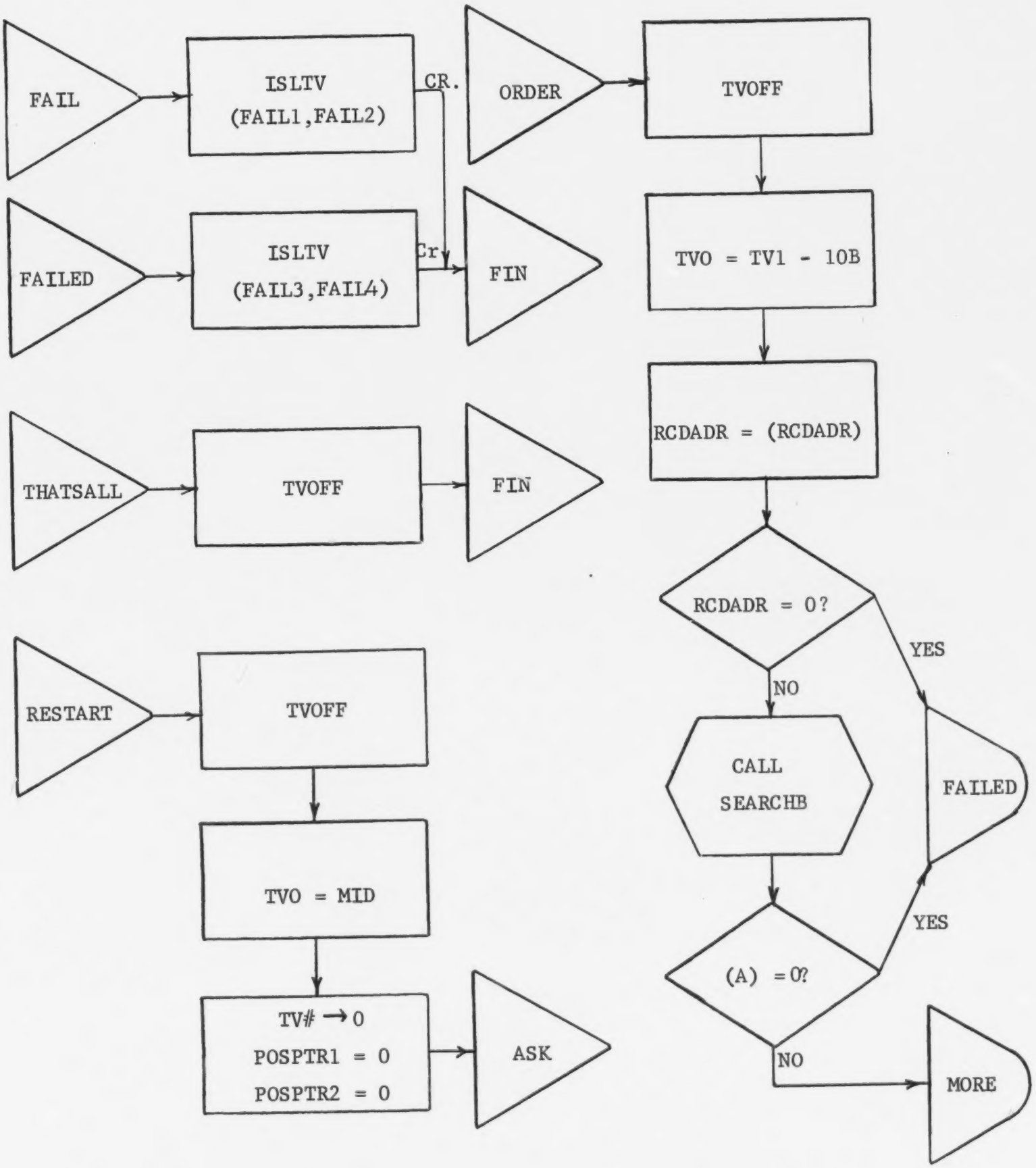
1. PTR1U
2. PTR1D
3. PTR2U
4. PTR2D
5. DELETE
6. ADD
7. UPTREE
8. DWNTREE
9. RESTART
10. STOR
11. LLINK
12. LISTALL
13. PRT
14. THATSALL
15. THATSALL

MAIN PROGRAM CONTINUE

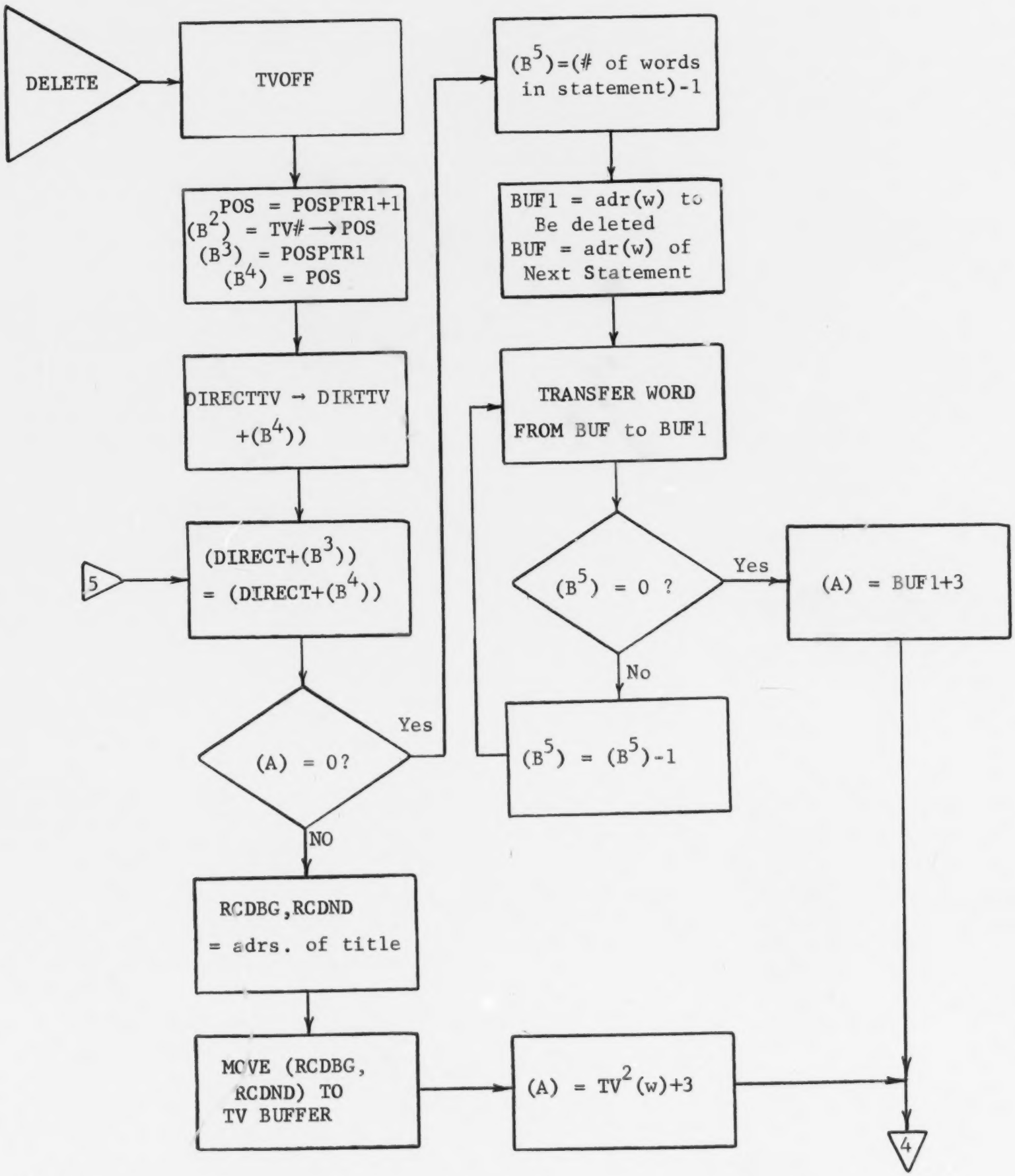




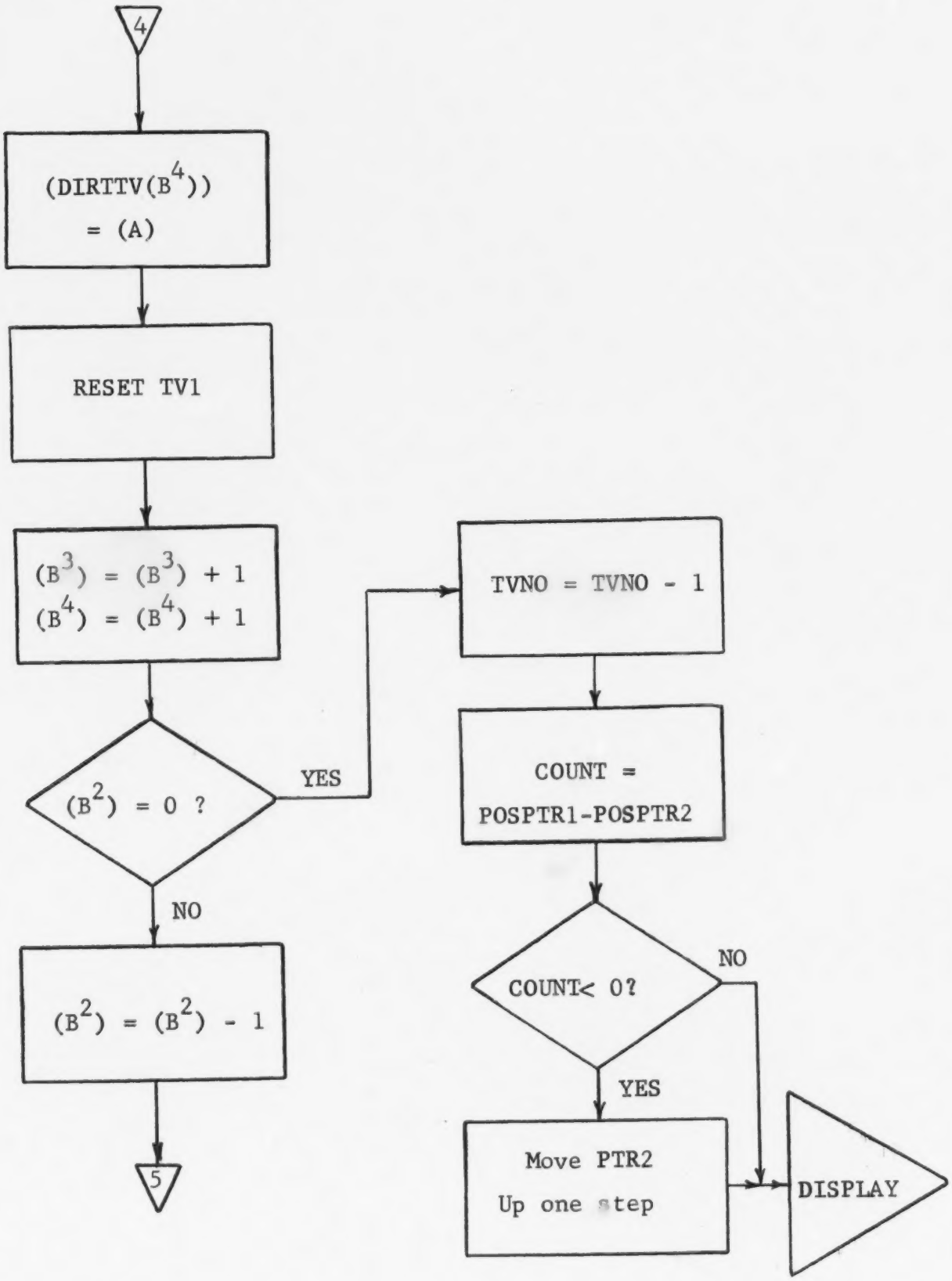
MAIN PROGRAM CONTINUE



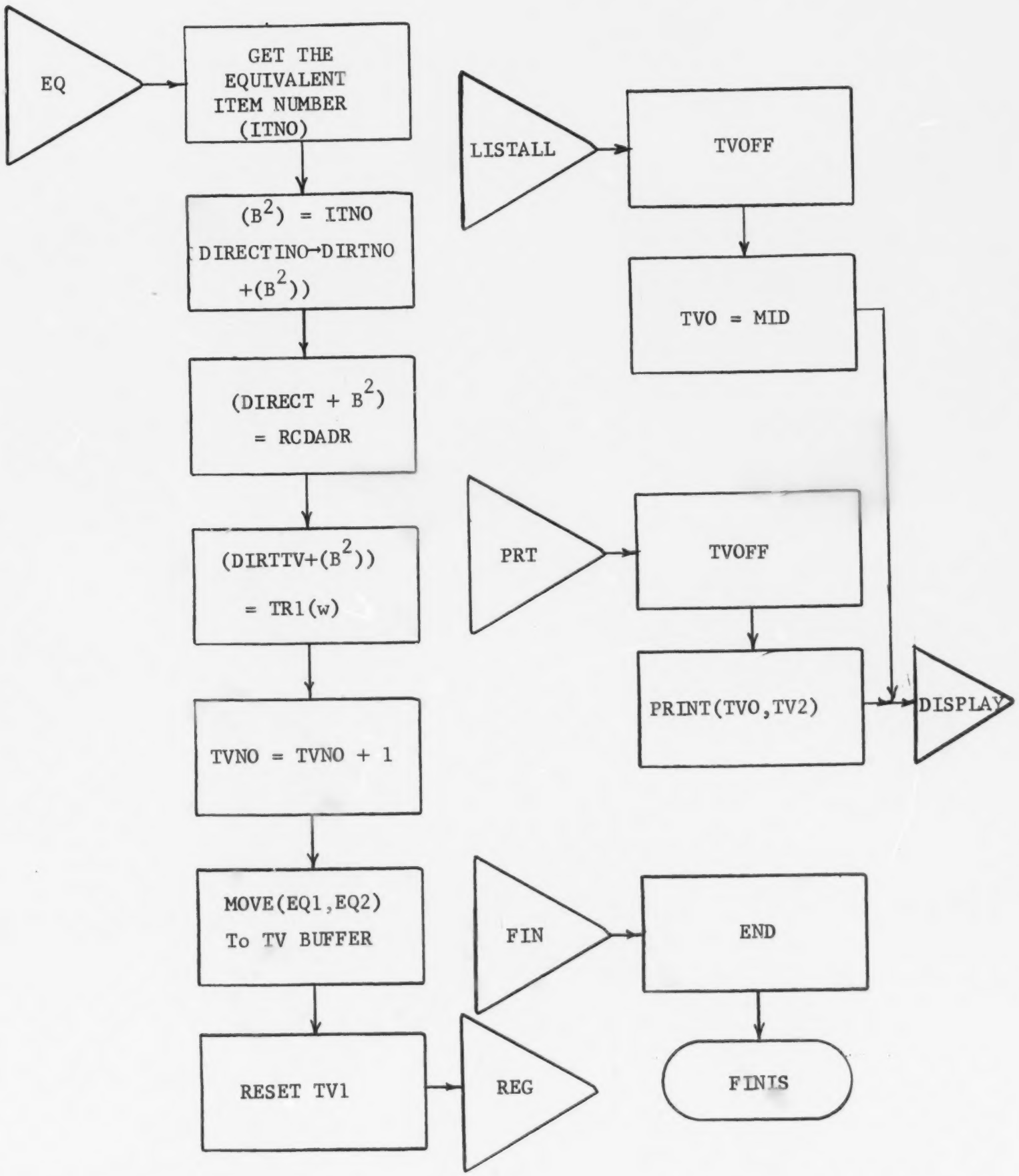
MAIN PROGRAM CONTINUE

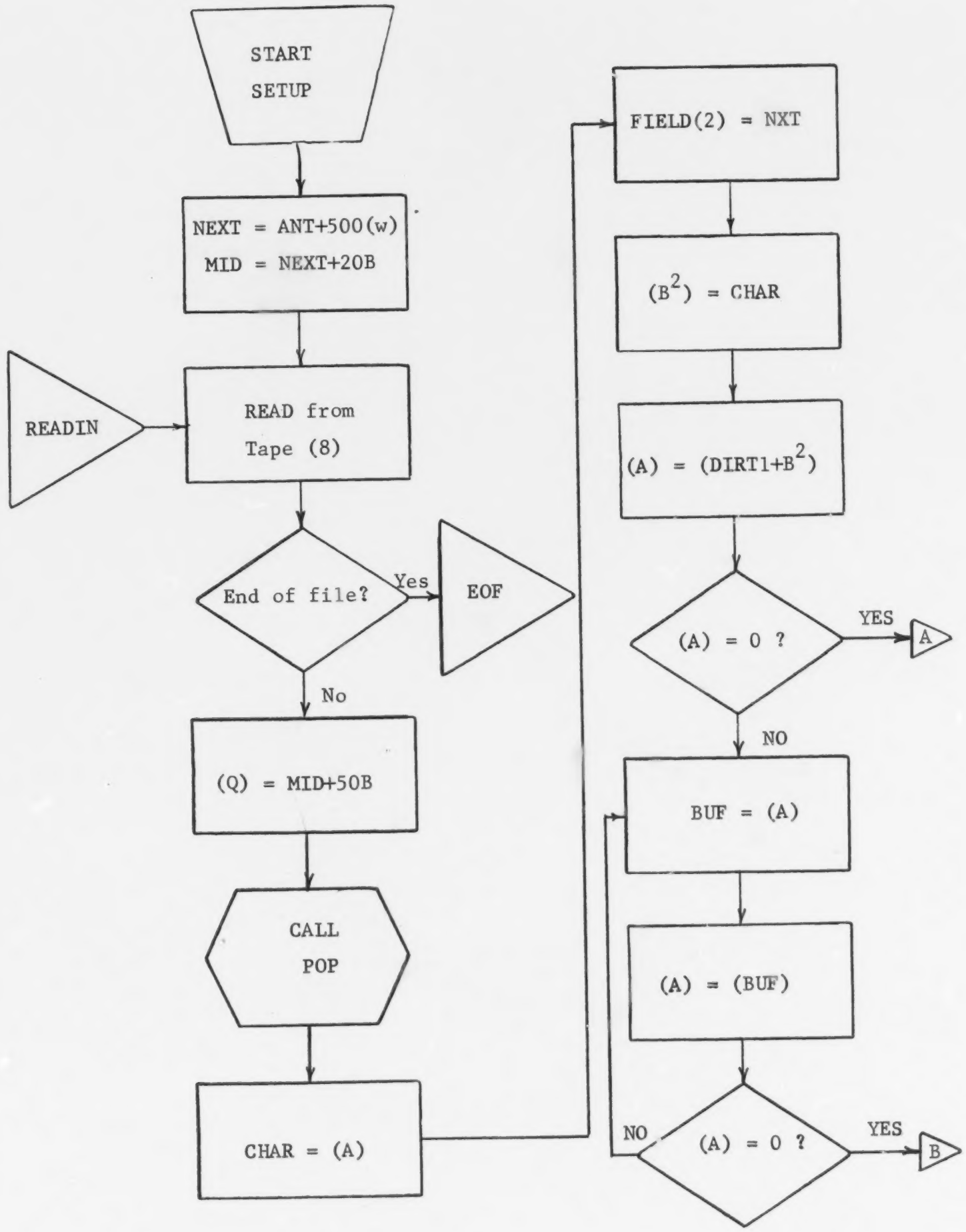


MAIN PROGRAM CONTINUE

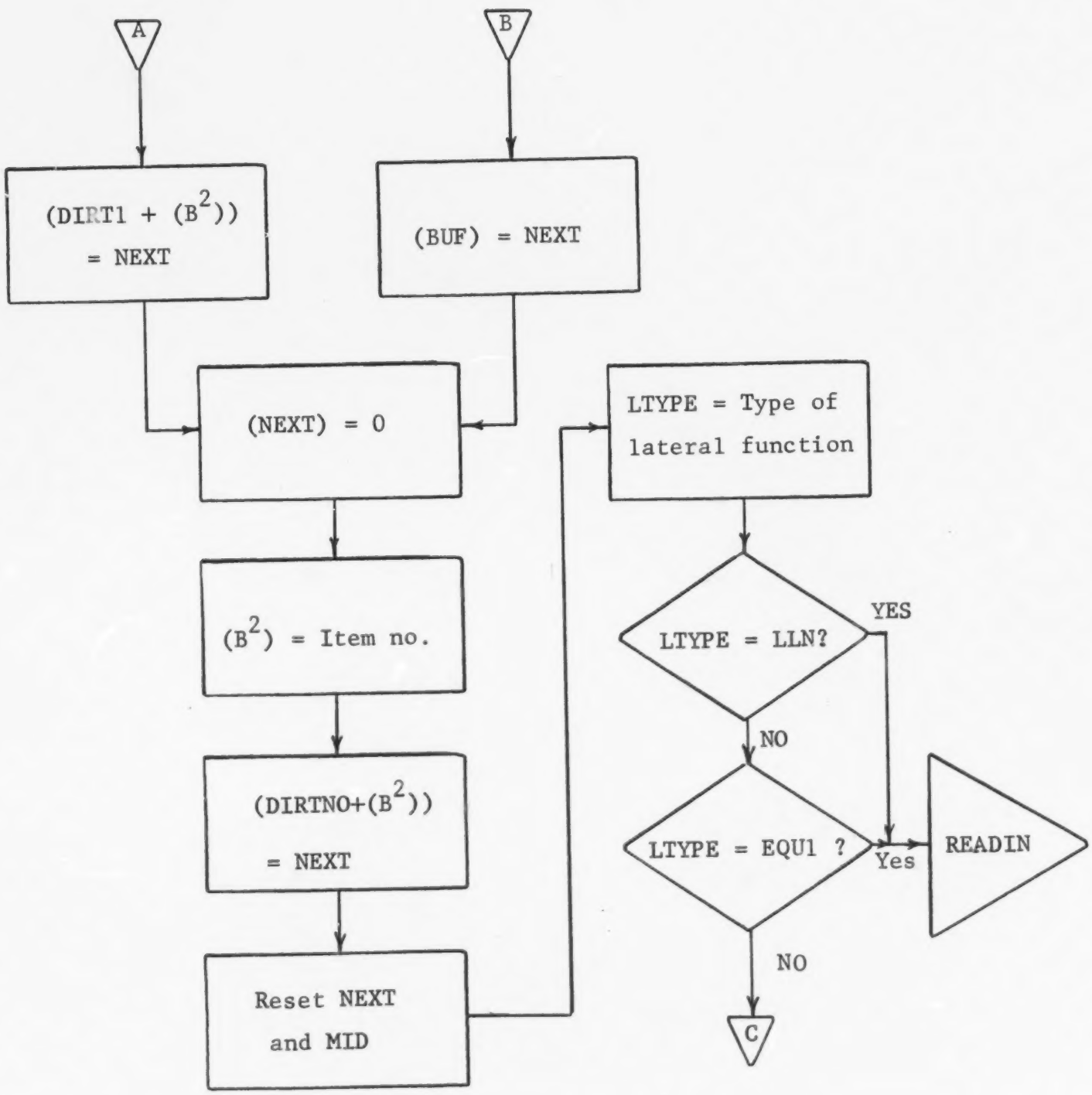


MAIN PROGRAM CONTINUE

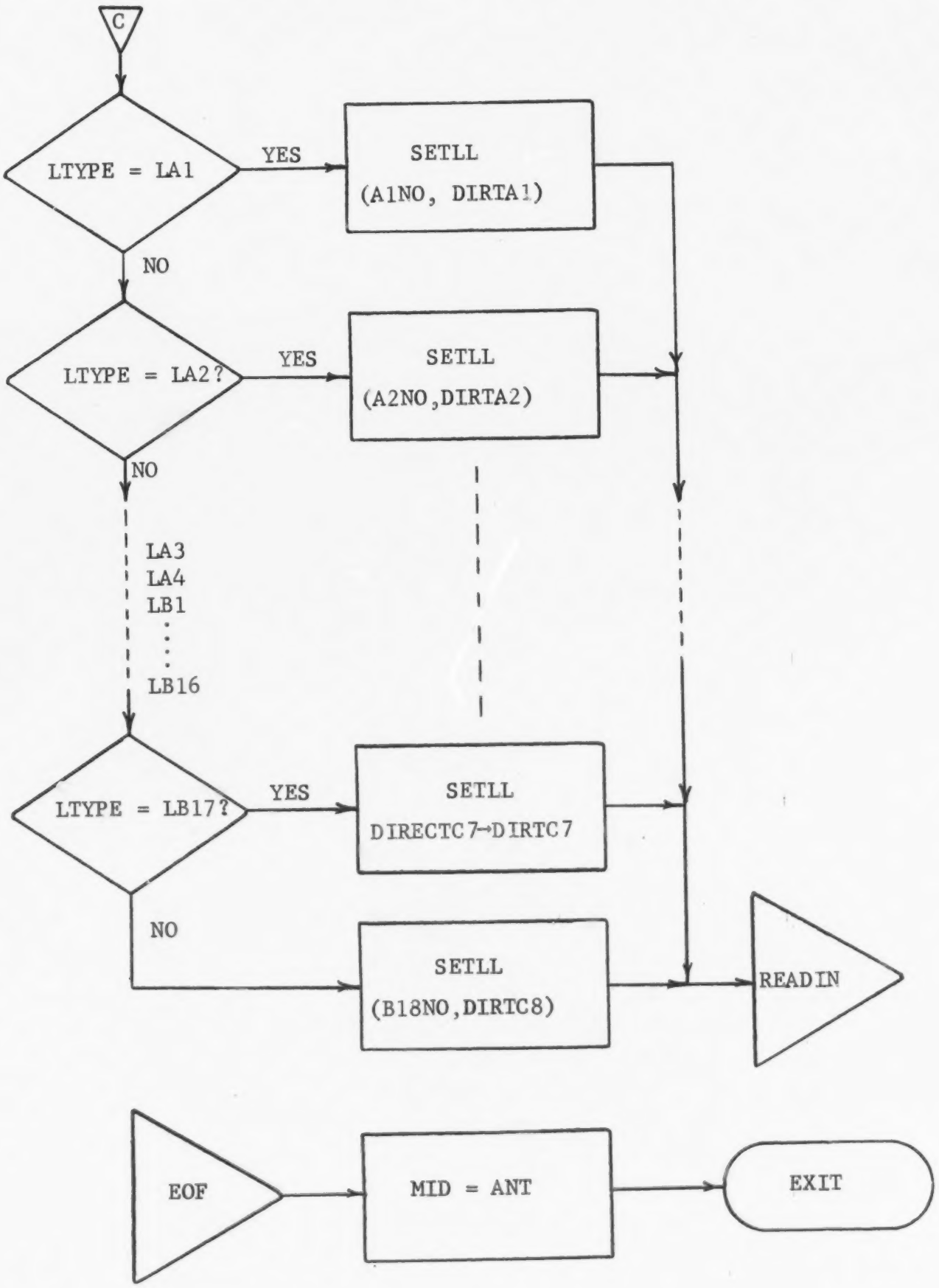


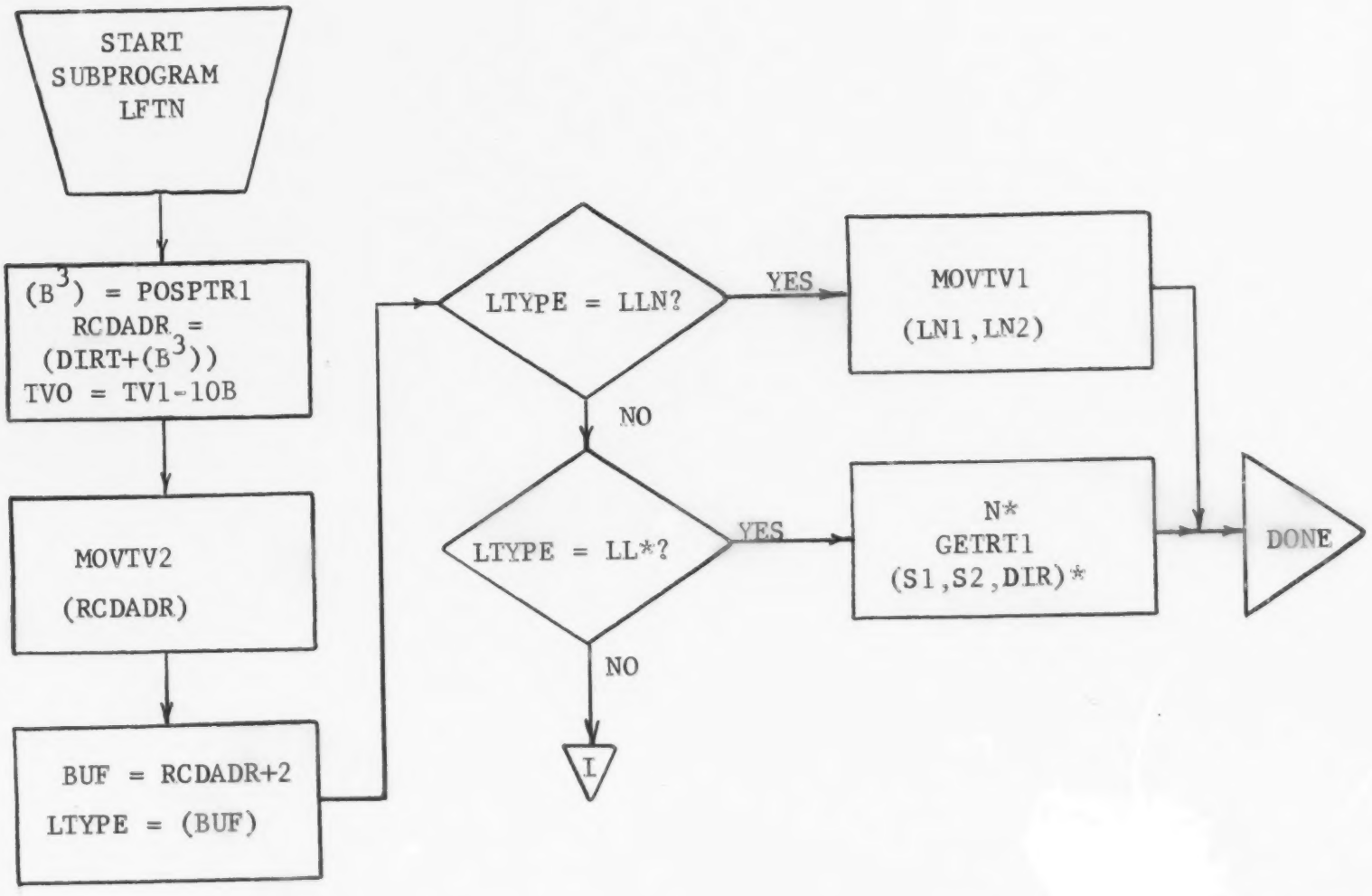

START
SETUP
NEXT = ANT+500(w)
MID = NEXT+20B
READIN
READ from
Tape (8)
End of file?
Yes
EOF
No
(Q) = MID+50B
CALL
POP
CHAR = (A)
FIELD(2) = NXT
(B2) = CHAR
(A) = (DIRT1+B2)
(A) = 0 ?
YES
A
NO
BUF = (A)
(A) = (BUF)
NO
(A) = 0 ?
YES
B

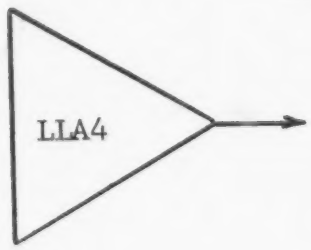
SUBPROGRAM SETUP CONTINUE



SUBPROGRAM SETUP CONTINUE

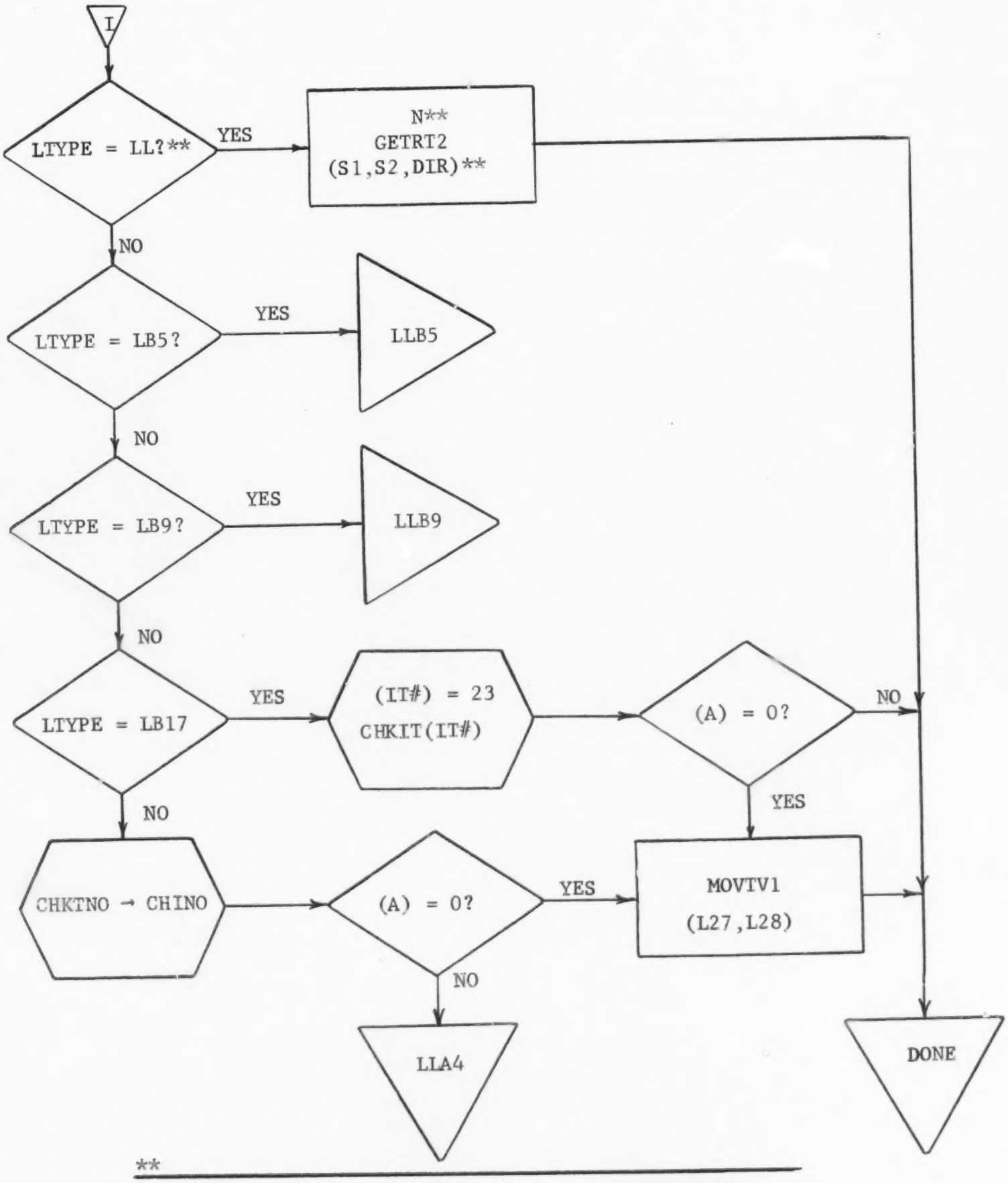






*

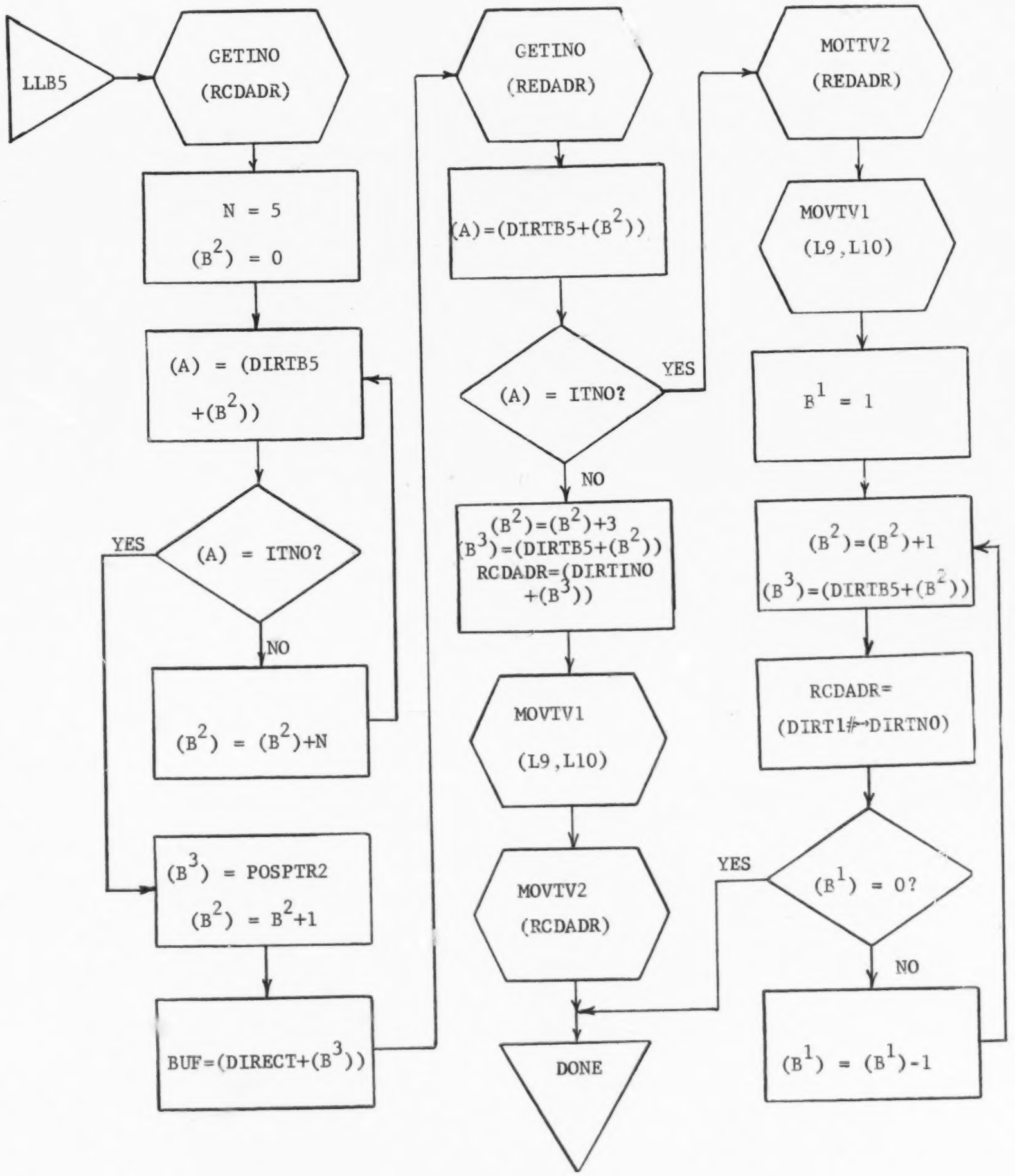
LL =	N =	(S1, S2, DIR) =
LA1	2	(L1, L2, DIRTA1)
LA2	2	(L3, L4, DIRTA2)
LA3	5	(L3, L4, DIRTA3)
LA4	3	(L3, L4, DIRTA4)
LB1	2	(L5, L6, DIRTB1)
LB2	3	(L5, L6, DIRTB2)
LB3	2	(L7, L8, DIRTB3)
LB4	3	(L7, L8, DIRTB4)
LB6	2	(L9, L10,DIRTB6)
LB7	3	(L9, L10,DIRTB7)
LB8	3	(L11,L12,DIRTB8)
LB10	3	(L15,L16,DIRTC0)
LB11	7	(L17,L18,DIRTC1)
LB12	8	(L17,L18,DIRTC2)
LB14	2	(L21,L22,DIRTC4)



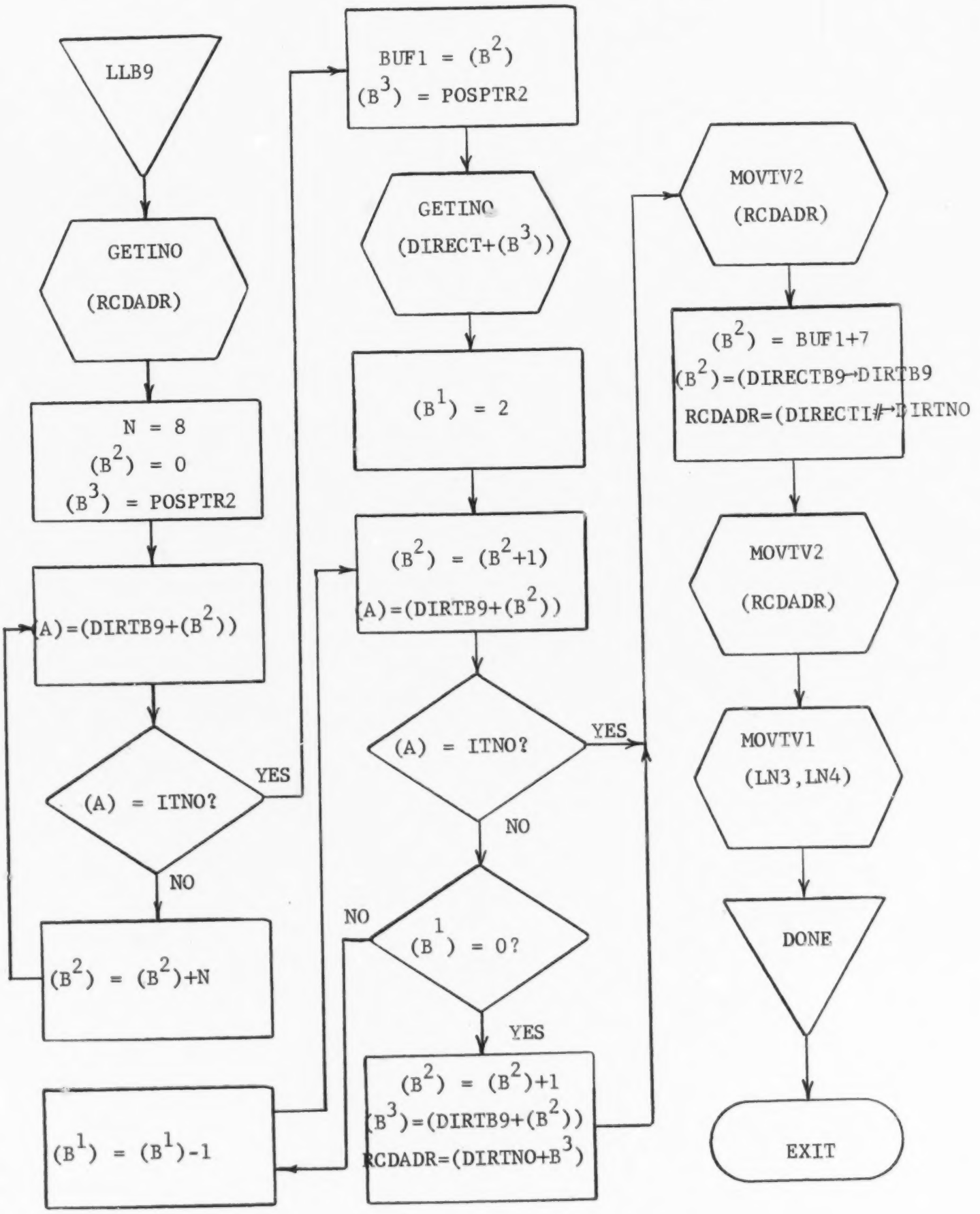
**

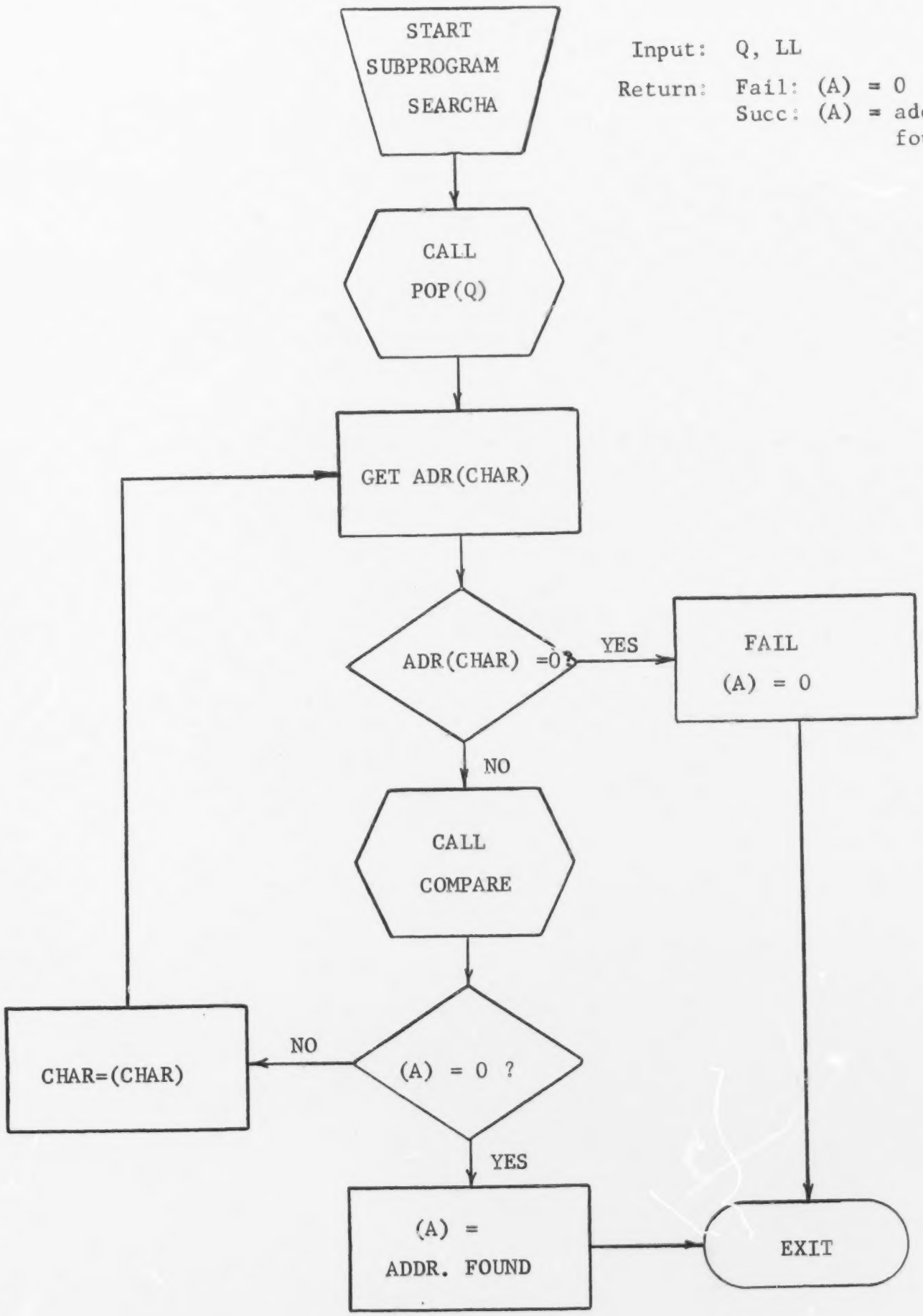
LL =	N =	(S1,S2,DIR)
LB13	3	(L19, L20, DIRTC3)
LB15	3	(L23, L24, DIRTC5)
LB16	3	(L25, L26, DIRTC6)

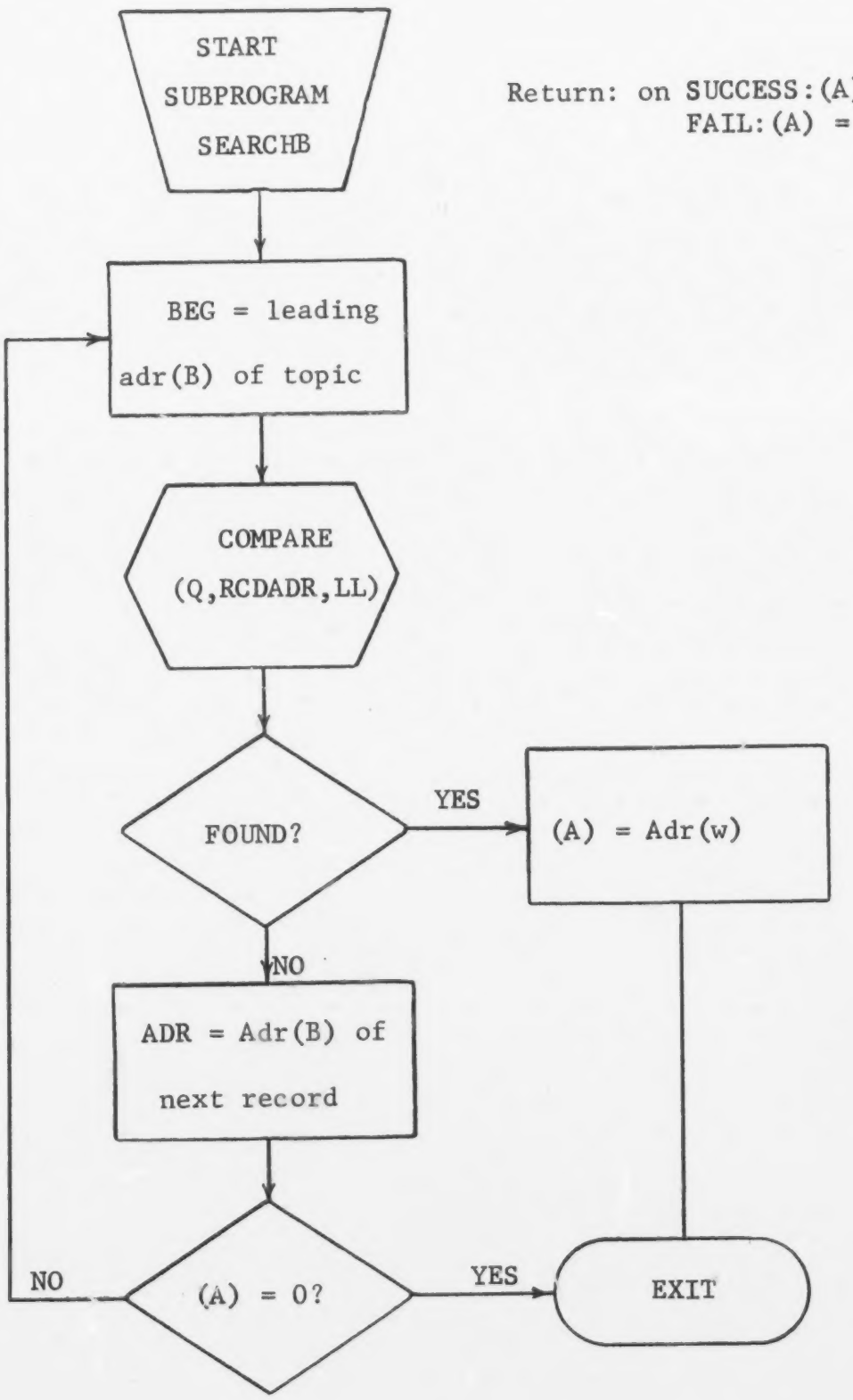
SUBPROGRAM LFTN CONTINUE



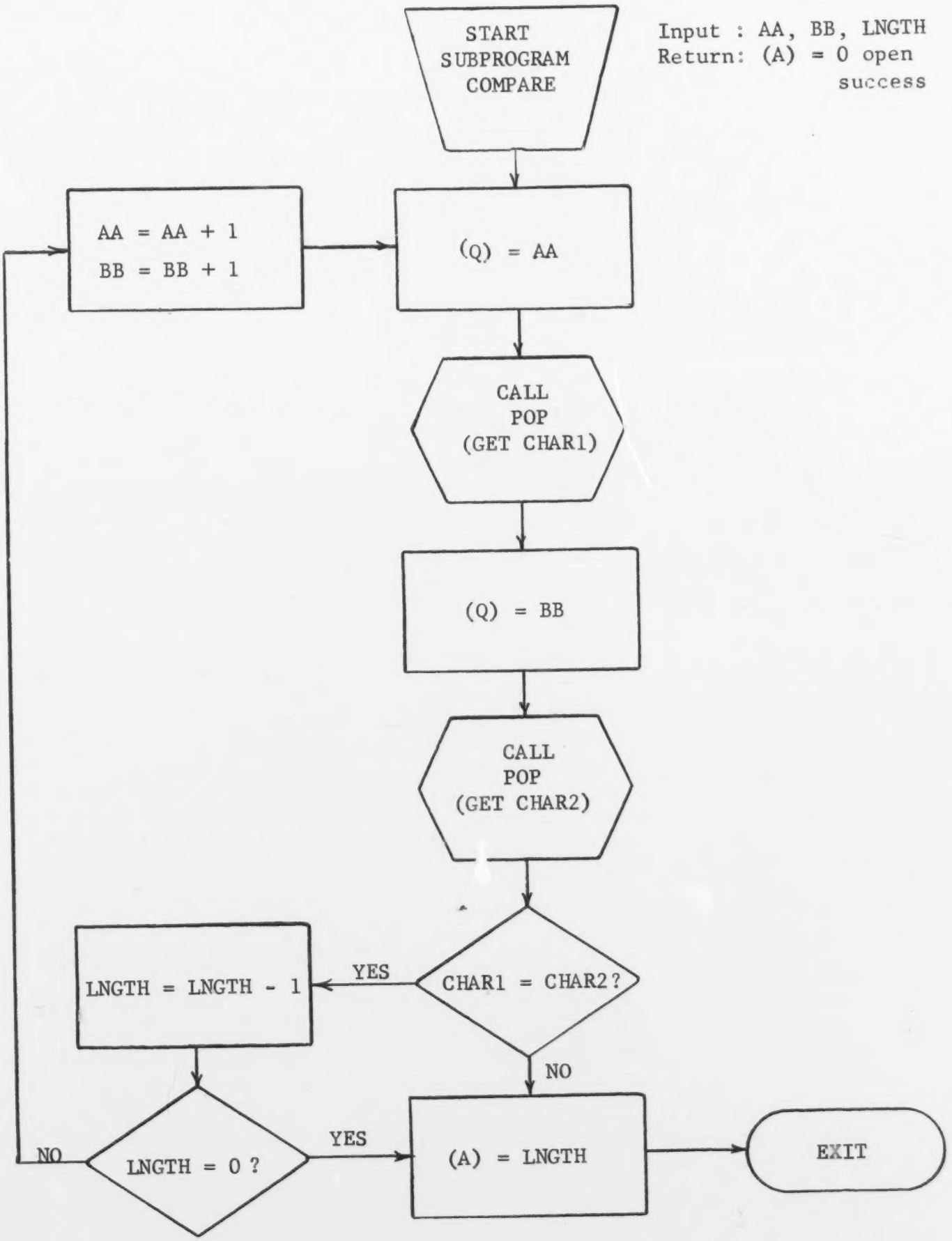
SUBPROGRAM LFTN CONTINUE

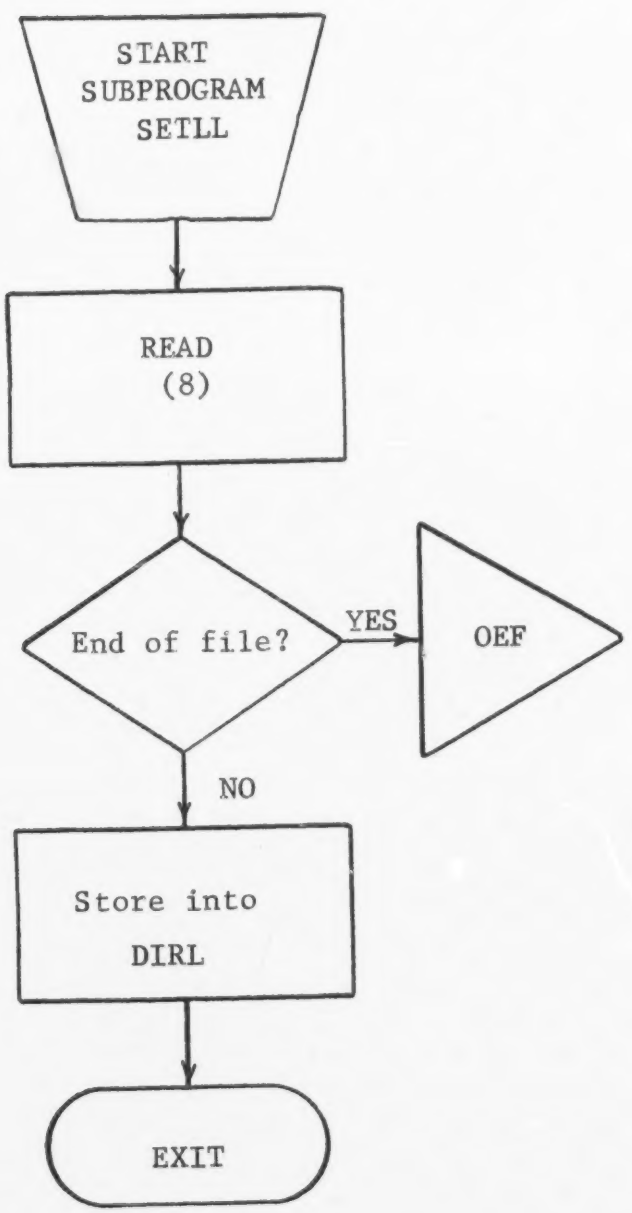


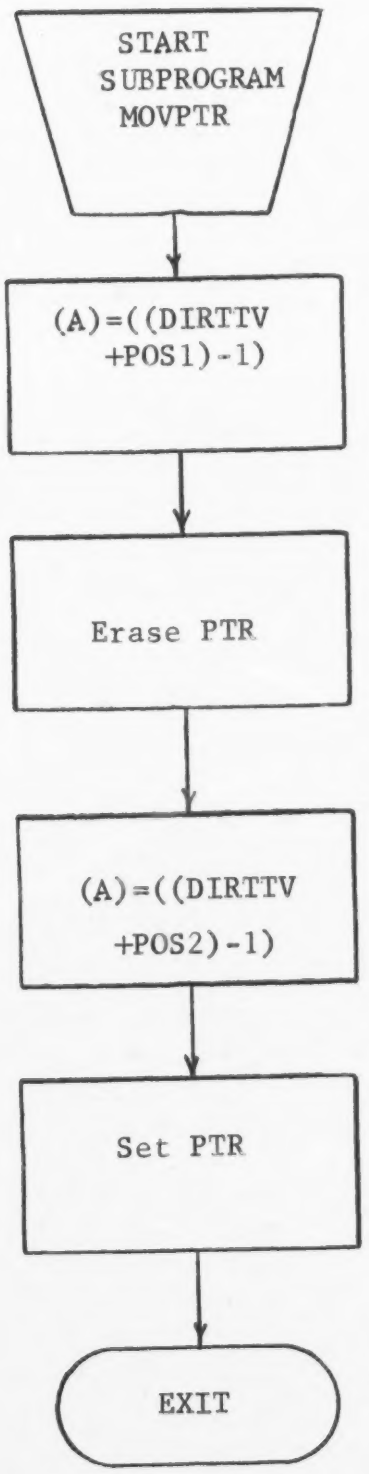

START
SUBPROGRAM
SEARCHA
Input: Q, LL
Return: Fail: (A) = 0
Succ: (A) = addr. found
CALL
POP(Q)
GET ADR(CHAR)
ADR(CHAR) =0?
YES
FAIL
(A) = 0
NO
CALL
COMPARE
(A) = 0 ?
NO
CHAR=(CHAR)
YES
(A) =
ADDR. FOUND
EXIT

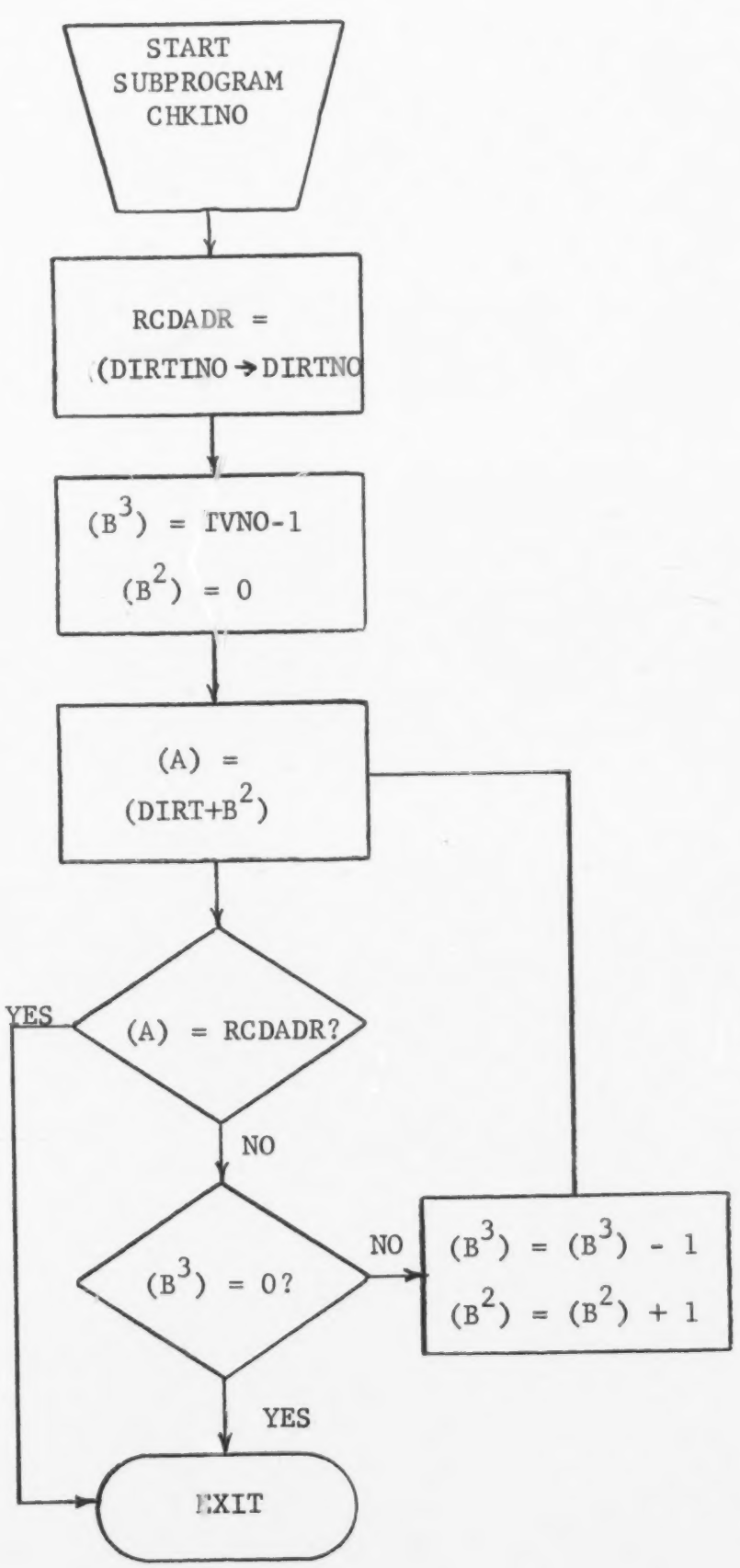


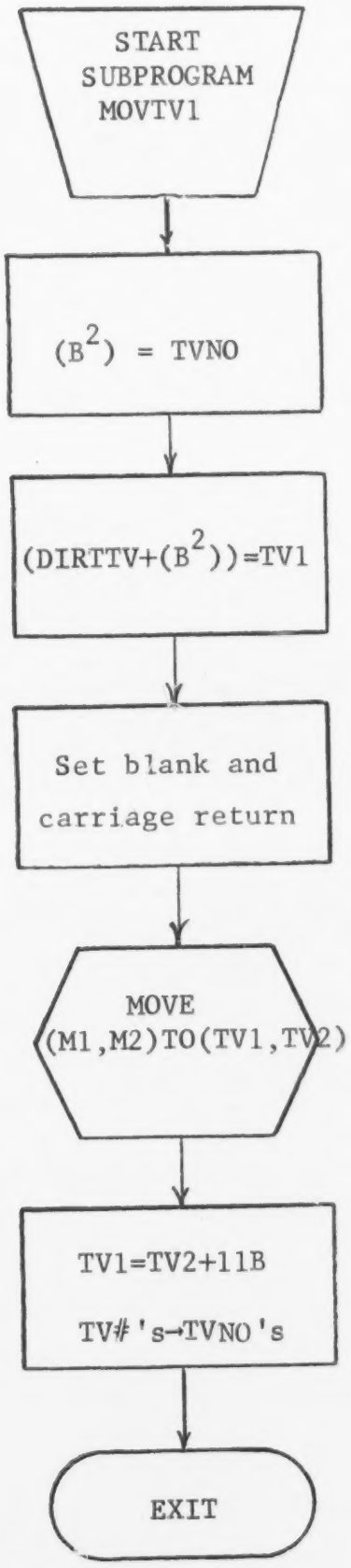
Return: on SUCCESS: (A) = Adr(w)
FAIL: (A) = 0

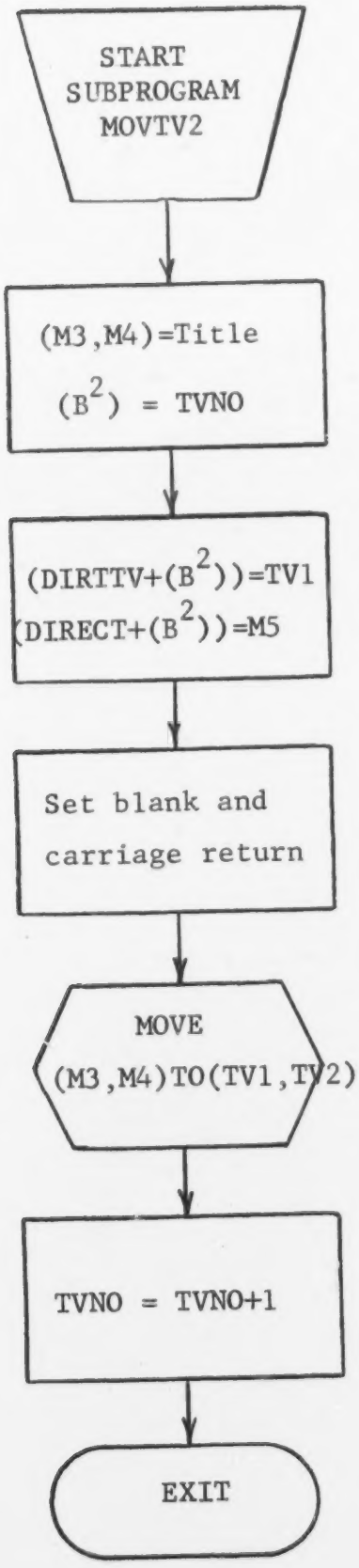

START
SUBPROGRAM
COMPARE
Input : AA, BB, LNGTH
Return: (A) = 0 open
success
AA = AA + 1
BB = BB + 1
(Q) = AA
CALL
POP
(GET CHAR1)
(Q) = BB
CALL
POP
(GET CHAR2)
CHAR1 = CHAR2?
YES
LNGTH = LNGTH - 1
NO
LNGTH = 0 ?
NO
YES
(A) = LNGTH
EXIT

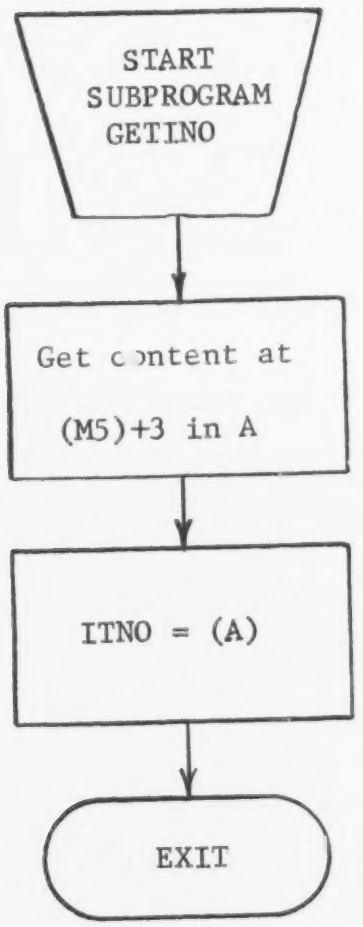

START
SUBPROGRAM
SETLL
READ
(8)
End of file?
YES
OEF
NO
Store into
DIRL
EXIT


START
SUBPROGRAM
MOVPTR
(A)=((DIRTTV
+POS1)-1)
Erase PTR
(A)=((DIRTTV
+POS2)-1)
Set PTR
EXIT


START
SUBPROGRAM
CHKINO
RCDADR =
(DIRTINO → DIRTNO
(B3) = TVNO-1
(B2) = 0
(A) =
(DIRT+B2)
(A) = RCDADR?
YES
NO
(B3) = 0?
NO
(B3) = (B3) - 1
(B2) = (B2) + 1
YES
EXIT


START
SUBPROGRAM
MOVTV1
(B2) = TVNO
(DIRTTV+(B2))=TV1
Set blank and
carriage return
MOVE
(M1,M2)TO(TV1,TV2)
TV1=TV2+11B
TV#'s→TVNO's
EXIT

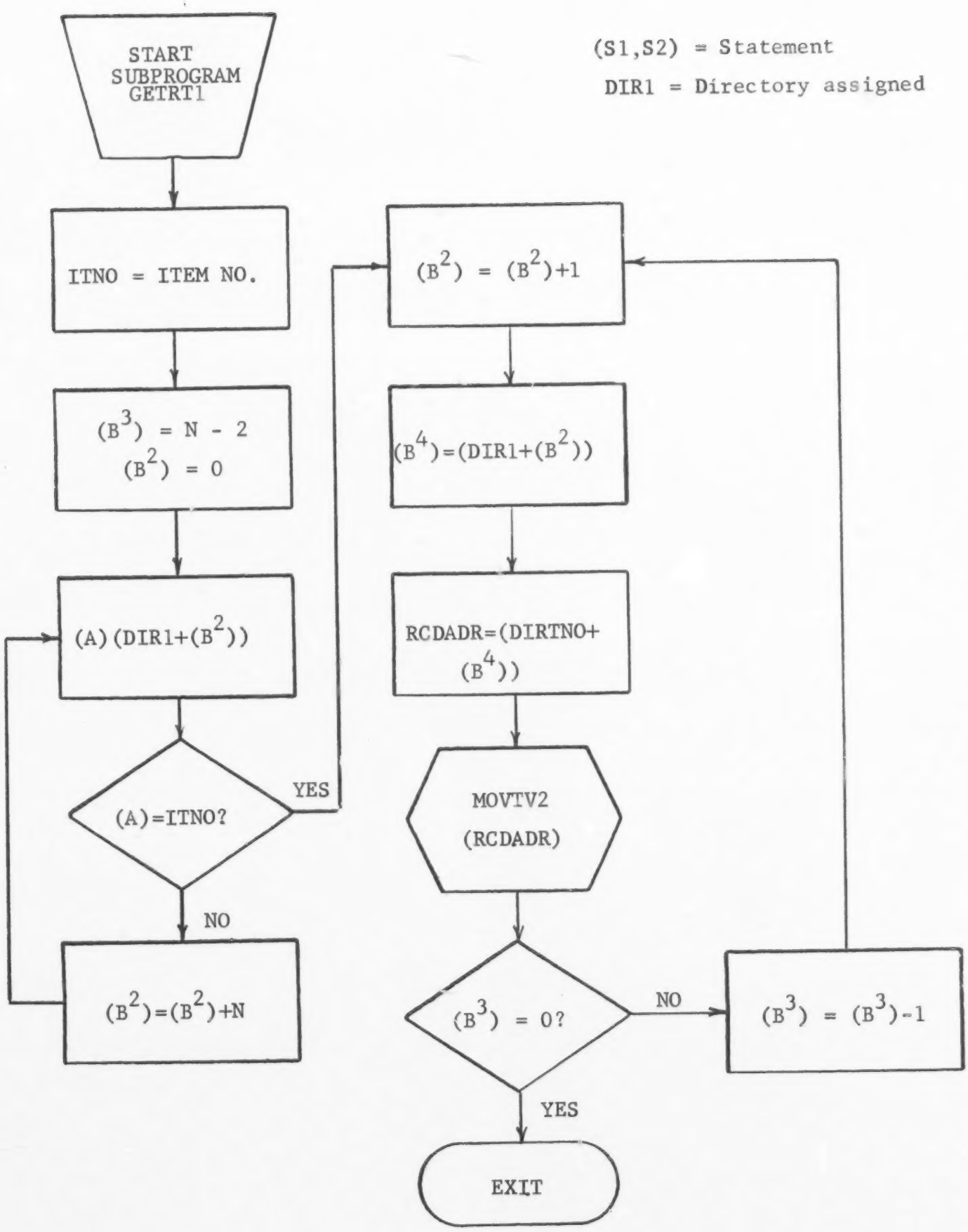


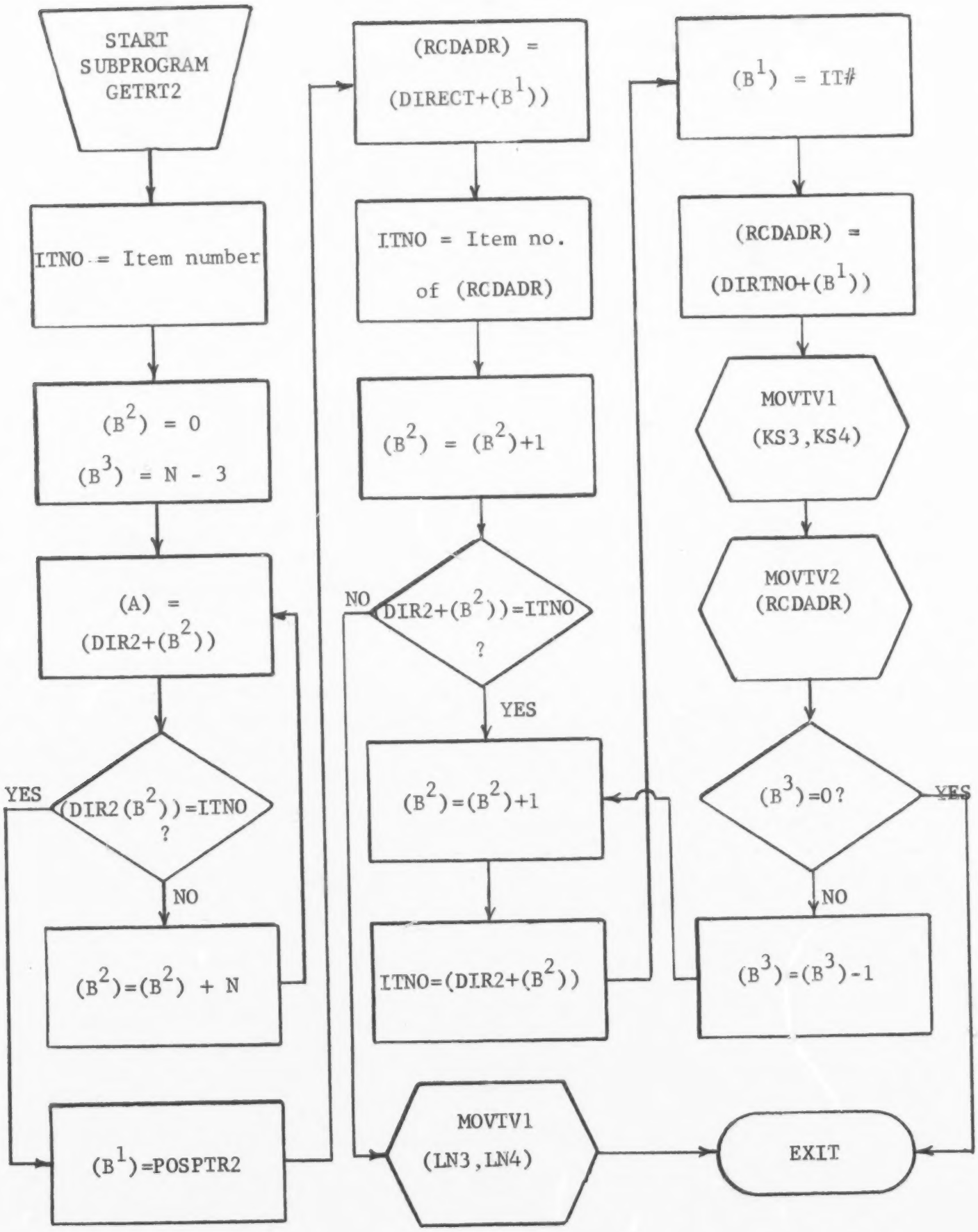


M5 = leading address
of record.

(S1,S2) = Statement

DIR1 = Directory assigned




START SUBPROGRAM GETRT2
ITNO = Item number
(B2) = 0
(B3) = N - 3
(A) = (DIR2+(B2))
YES
(DIR2 (B2)) = ITNO ?
NO
(B2) = (B2) + N
(B1) = POSPTR2
(RCDADR) = (DIRECT+(B1))
ITNO = Item no. of (RCDADR)
(B2) = (B2)+1
NO
DIR2+(B2)) = ITNO ?
YES
(B2) = (B2)+1
ITNO=(DIR2+(B2))
MOVTV1 (LN3, LN4)
(B1) = IT#
(RCDADR) = (DIRTNO+(B1))
MOVTV1 (KS3, KS4)
MOVTV2 (RCDADR)
(B3) = 0 ?
YES
NO
(B3) = (B3) - 1
EXIT

Distribution List as of August 1, 1970

ESD (ESTI)
L. G. Hanscom Field
Bedford, Mass 01731 2 Copies

Mr I. A. Balton
Institute for Exploratory Research
Code: AMSEL-XL
U. S. Army Electronics Command
Fort Monmouth, New Jersey 07703

LTC Howard W. Jackson
Deputy Dir of Electr. & Solid St. Sciences
Air Force Office of Scientific Research
1400 Wilson Boulevard
Arlington, Virginia 22209 5 Copies

Defense Documentation Center
Attn: DDC-TCA
Cameron Station
Alexandria, Virginia 22314 50 Copies

Director, Electronics Program
Attn: Code 427
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217 3 Copies

Naval Air Systems Command
AIR 03
Washington, D. C. 20360 2 Copies

Naval Electronic Systems Command
ELEX 03, Room 2046 Munitions Building
Department of the Navy
Washington, D.C. 20360 2 Copies

Director
Naval Research Laboratory
Washington, D.C. 20390
Attn: Code 2027 6 Copies

Commander
U. S. Naval Ordnance Laboratory
Attn: Librarian
White Oak, Md. 20910 2 Copies

Commanding General
Attn: STEWS-RE-L, Technical Library
White Sands Missile Range
New Mexico 88002 2 Copies

Commander
Naval Electronics Laboratory Center
Attn: Library
San Diego, Calif 92152 2 Copies

Dr L. M. Hollingsworth
AFCRL (CRN)
L. G. Hanscom Field
Bedford, Massachusetts 01731

Division of Engineering & Applied Physics
210 Pierce Hall
Harvard University
Cambridge, Massachusetts, 02138

Director
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Miss R. Joyce Harman
Project MAC, Room 810
545 Technology Square
Cambridge, Mass 02139

Professor R. H. Rediker
Electrical Engineering Prof.
Mass. Institute of Technology
Building 13-3050
Cambridge, Mass 02139

Raytheon Company
Research Division Library
28 Seyon Street
Waltham, Massachusetts 02154

Sylvania Electronic Systems
Applied Research Laboratory
Attn: Documents Librarian
40 Sylvan Road
Waltham, Mass 02154

Commanding Officer
Army Materials & Mechanics Res. Center
Attn: Dr H. Priest
Watertown Arsenal
Watertown, Massachusetts 02172

MIT Lincoln Laboratory
Attn: Library A-082
P.O. Box 73
Lexington, Mass. 02173

Commanding Officer
Office of Naval Research Branch Office
495 Summer Street
Boston, Massachusetts 02210

Commanding Officer (Code 2064)
U. S. Naval Underwater Sound Laboratory
Fort Trumbull
New London, Connecticut 06320

Dept of Eng & Applied Science
Yale University
New Haven, Conn 06520

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-CT-A
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-CT-D
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-CT-I
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-CT-L (Dr W. S. McAfee)
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-CT-O
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-CT-R
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Fort Monmouth, New Jersey 07703
Attn: AMSEL-CT-S

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-DL
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-GG-DD
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-KL-D
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-KL-E
Fort Monmouth, New Jersey 07707

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-KL-I
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-KL-SM
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-KL-S
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-KL-T
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-NL-A
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-NL-C
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-NL-D (Dr H. Bennett)
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-NL-P
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-SC
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-VL-D
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-VL-F
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-WL-D
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-XL-DT
Fort Monmouth, New Jersey 07703

Commanding General
U. S. Army Electronics Command
Attn: AMSEL-XL-D
Fort Monmouth, New Jersey 07703

Mr Norman J. Field, AMSEL-RD-S
Chief, Office of Science & Technology
Research and Development Directorate
U. S. Army Electronics Command
Fort Monmouth, New Jersey 07703

Project Manager
Common Positioning & Navigation Systems
Attn: Harold H. Bahr (AMCPM-NS-TM),
Bldg. 439
U. S. Army Electronics Command
Fort Monmouth, New Jersey 07703

U. S. Army Munitions Command
Attn: Science & Technology
Info Br., Bldg 59
Picatinny Arsenal, SMUPA-RT-S
Dover, New Jersey 07801

European Office of Aerospace Research
Technical Information Office
Box 14, FPO New York 09510

Director
Columbia Radiation Laboratory
Columbia University
538 West 120th St.
New York, N.Y. 10027

New York University
Engineering Library
Bronx, New York 10453

Mr Jerome Fox, Research Coordinator
Polytechnic Institute of Brooklyn
333 Jay St.
Brooklyn, N. Y. 11201

Airborne Instruments Laboratory
Deerpark, New York 11729

Dr. W. R. Lepage, Chairman
Syracuse University
Dept of Electrical Engineering
Syracuse, N. Y. 13210

Rome Air Development Center
Attn: Documents Library (EMTLD)
Griffiss Air Force Base, N. Y. 13440

Mr H. E. Webb (EMBIS)
Rome Air Development Center
Griffiss Air Force Base, N.Y. 13440

Professor James A. Cadzow
Department of Electrical Engineering
State University of New York at Buffalo
Buffalo, N. Y. 14214

Dr. A. G. Jordan
Head of Dept of Elec Engineering
Carnegie-Mellon University
Pittsburgh, Penn 15213

Hunt Library
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA. 15213

Lehigh University
Dept of Electrical Engineering
Bethelehem, Pennsylvania 18015

Commander (ADL)
Naval Air Development Center
Attn: NADC Library
Johnsville, Warminister, Pa 18974

Technical Director (SMUFA-A2000-107-1)
Frankford Arsenal
Philadelphia, Pennsylvania 19137

Mr M. Zane Thornton, Chief, Network
Engineering, Communications and
Operations Branch, Lister Hill
National Center/ Biomedical Communications
8600 Rockville Pike
Bethesda, Maryland 20014

U. S. Post Office Dept
Library- Room 6012
12th & Pennsylvania Ave. N.W.
Washington, D.C. 20260

Technical Library
DDR&E
Room 3C-122, The Pentagon
Washington, D.C. 20301

Distribution List (Cont'd.)

Director for Materials Sciences
Advanced Research Projects Agency
Department of Defense
Washington, D.C. 20301

Assistant Director, (Research)
Office of Director of Defense Research & Engineering
Pentagon, Rm 3C128
Washington, D.C. 20301

Chief, R & D Division (340)
Defense Communications Agency
Washington, D.C. 20305

Commanding General
U. S. Army Materiel Command
Attn: AMCRD-TP
Washington, D.C. 20315

Director, U. S. Army Material Concepts Agency
Washington, D.C. 20315

Hq USAF (AFRDD)
The Pentagon
Washington, D.C. 20330

Hq USAF (AFRDDG)
The Pentagon
Washington, D.C. 02330

Hq USAF (AFRDSD)
The Pentagon
Washington, D.C. 20330

AFSC (SCTSE)
Andrews Air Force Base, Maryland 20331

Dr I. R. Mirman
Hq AFSC (SGGP)
Andrews AFB, Maryland 20331

Naval Ship Systems Command
Ship 031
Washington, D.C. 20360

Naval Ship Systems Command
Ship 035
Washington, D.C. 20360

Commander
U. S. Naval Security Group Command
Attn: G43
3801 Nebraska Avenue
Washington, D.C. 20390

U. S. Naval Oceanographic Office
Attn: M. Rogofsky, Librarian (Code 640)
Washington, D.C. 20390

Director
Naval Research Laboratory
Washington D.C. 20390
Attn: Dr A. Brodizinsky, Sup. Elec Div

Director
Naval Research Laboratory
Washington, D.C. 20390
Attn: Maury Center Library (Code 8050)

Director
Naval Research Laboratory
Washington, D.C. 20390
Attn: Dr W. C. Hall, Code 7000

Director
Naval Research Laboratory
Attn: Library, Code 2029 (ONRL)
Washington, D.C. 20390

Dr G.M.R. Winkler
Director, Time Service Division
U. S. Naval Observatory
Washington, D.C. 20390

Colonel E. P. Gaines, Jr
ACDA/FO
1901 Pennsylvania Ave. N. W.
Washington, D.C. 20451

Commanding Officer
Harry Diamond Laboratories
Attn: Mr Berthold Altman (AMXD)-TI)
Connecticut Ave & Van Ness St., N.W.
Washington, D.C. 20438

Central Intelligence Agency
Attn: CRS/ADD Publications
Washington, D.C. 20505

Dr H. Harrison, Code RRE
Chief, Electrophysics Branch
National Aeronautics & Space Admin.
Washington, D.C. 20546

The John Hopkins University
Applied Physics Laboratory
Attn: Document Librarian
8621 Georgia Avenue
Silver Spring, Maryland 20910

Commanding Officer (AMXRD-BAT)
U.S. Army Ballistics Research Laboratory
Aberdeen Proving Ground
Aberdeen, Maryland 21005

Technical Director
U. S. Army Land Warfare Laboratory
Aberdeen Proving Ground
Aberdeen, Maryland 21005

Electromagnetic Compatibility
Analysis Center (ECAC)
Attn: (ACOAT)
North Severn
Annapolis, Maryland 21402

Commanding Officer
U. S. Army Engineer Topographic Labs
Attn: STINFLO Center
Fort Belvoir, Virginia 22060

Director (NV-D)
Night Vision Laboratory, USAECOM
Fort Belvoir, Virginia 22060

U. S. Army Mobility Equipment Research and Development Center
Attn: Technical Document Center Bldg 315
Fort Belvoir, Virginia 22060

Dr Alvin D. Schnitzler
Institute for Defense Analyses
Science and Technology Division
400 Army-Navy Drive
Arlington, Virginia 22202

Director, Physical & Eng. Sciences Div.
3045 Columbia Pike
Arlington, Virginia 22204

Commanding General
U. S. Army Security Agency
Attn: IARD-T
Arlington Hall Station
Arlington, Virginia 22212

Dr Joel Trimble, Code 437
Information Systems Branch
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Commanding General
USACDC Institute of Land Combat
Attn: Technical Library, rm 636
2461 Eisenhower Avenue
Alexandria, Virginia 22314

VELA Seismological Center
300 North Washington St.
Alexandria, Virginia 22314
U. S. Naval Weapons Laboratory
Dahlgren, Virginia 22448

Research Laboratories for the Eng. Sciences, School of Engineering & Applied Science
University of Virginia
Charlottesville, Va. 22903

Dr Herman Robl
Deputy Chief Scientist
U. S. Army Research Office (Durham)
Box CM, Duke Station
Durham, North Carolina 27706

Richard O. Ulsh (CRDARD-IP)
U. S. Army Research Office (Durham)
Box CM, Duke Station
Durham, North Carolina 27706

ADTC (ADBPS-12)
Eglin AFB, Florida 32542

Commanding Officer
Naval Training Device Center
Orlando, Florida 32813

Technical Library, AFETR
(ETV,MU-135)
Patrick AFB, Florida 32925

Commanding General
U. S. Army Missile Command
Attn: AMSMI-RR
Redstone Arsenal, Alabama 35809

Redstone Scientific Information Center
Attn: Chief, Document Section
U. S. Army Missile Command
Redstone Arsenal, Alabama 35809

AUL3T-9663
Maxwell AFB, Alabama 36112

Hq AEDC (AETS)
Attn: Library/Documents
Arnold AFS, Tennessee 37389

Case Institute of Technology
Engineering Division
University Circle
Cleveland, Ohio 44106

NASA Lewis Research Center
Attn: Library
21000 Brookpark Road
Cleveland, Ohio 44135

Director
Air Force Avionics Laboratory
Wright-Patterson AFB, Ohio 45433
AFAL (AVTA) R. D. Larson
Wright-Patterson AFB, Ohio 45433

AFAL (AVT) Dr H. V. Noble, Chief
Electronics Technology Division
Air Force Avionics Laboratory
Wright-Patterson AFB, Ohio 45433

Dr Robert E. Fontana
Head, Dept of Electrical Engineering
Air Force Institute of Technology
Wright Patterson AFB, Ohio 45433

Dept of Electrical Engineering
Clippinger Laboratory
Ohio University
Athens, Ohio 45701

Commanding Officer
Naval Avionics Facility
Indianapolis, Indiana 46241

Dr John C. Hancock, Head
School of Electrical Engineering
Purdue University
Lafayette, Ind 47907

Professor Joseph E. Rowe
Chairman, Dept of Electrical Engineering
The University of Michigan
Ann Arbor, Michigan 48104

Dr G. J. Murphy
The Technological Institute
Northwestern University
Evanston, Ill 60201

Commanding Officer
Office of Naval Research Branch Office
219 South Dearborn St.
Chicago, Illinois 60604

Illinois Institute of Technology
Dept of Electrical Engineering
Chicago, Illinois 60616

Deputy for Res. and Eng (AMSE-DRE)
U. S. Army Weapons Command
Rock Island Arsenal
Rock Island, Illinois 61201

Commandant
U. S. Army Command & General Staff College
Attn: Acquisitions, Library Division
Fort Leavenworth, Kansas 66027

Dept of Electrical Engineering
Rice University
Houston, Texas 77001

HQ AMD (AMR)
Brooks AFB, Texas 78235

USAFSAM (SMKOR)
Brooks AFB, Texas 78235

Mr B. R. Locke
Technical Adviser, Requirements
USAF Security Service
Kelly Air Force Base, Texas 78241

Director Electronics Research Center
The University of Texas at Austin
Eng-Science Bldg 110
Austin, Texas 78712

Department of Electrical Engineering
Texas Technological University
Lubbock, Texas 79409

Commandant
U. S. Army Air Defense School
Attn: Missile Sciences Div., C&S Dept
P. O. Box 9390
Fort Bliss, Texas 79916

Director
Aerospace Mechanics Sciences
Frank J. Seiler Research Laboratory (OAR)
USAF Academy
Colorado Springs, Colorado 80840

Director of Faculty Research
Department of the Air Force
U. S. Air Force Academy
Colorado Springs, Colorado 80840

Major Richard J. Gowen
Tenure Associate Professor
Dept of Electrical Engineering
U. S. Air Force Academy
Colorado Springs, Colorado 80840

Distribution List (Cont'd.)

Academy Library (DFSLB)
U. S. Air Force Academy
Colorado Springs, Colorado 80840

M. A. Rothenberg (STEPD-SC(S))
Scientific Director
Desert Test Center
Bldg. 100, Soldiers Circle
Fort Douglas, Utah 84113

Utah State University
Dept of Electrical Engineering
Logan, Utah 84321

School of Engineering Sciences
Arizona State University
Tempe, Ariz 85281

Commanding General
U. S. Army Strategic Communications
Command
Attn: SCC-CG-SAE
Fort Huachuca, Arizona 85613

The University of Arizona
Dept of Electrical Engineering
Tucson, Arizona 85721

Cpt C. E. Baum
AFWL (WLRE)
Kirkland AFB, New Mexico 78117

Los Alamos Scientific Laboratory
Attn: Report Library
P. O. Box 1663
Los Alamos, N.M 87544

Commanding Officer
Atmospheric Sciences Laboratory
White Sands Missile Range, N. Mex 88002

Commanding Officer
(AMSEL-BL-WS-R)
Atmospheric Sciences Laboratory
White Sands Missile Range
New Mexico 88002

Chief, Missile Electronic Warfare
Tech Area
(AMSEL-WL-M)
Electronic Warfare Laboratory, USACOM
White Sands Missile Range, N.M. 88002

Director
Electronics Sciences Lab
University of Southern California
Los Angeles, Calif 90007

Engineering & Mathematical Sciences Library
University of California at Los Angeles
405 Hilgred Avenue
Los Angeles, Calif. 90024

Aerospace Corporation
P. O. Box 95085
Los Angeles, California 90045
Attn: Library Acquisitions Group

Hq SAMSO (SMTTA/Lt Belate)
AF Unit Post Office
Los Angeles, Calif. 90045

Dr Sheldon J. Wells
Electronic Properties Information Center
Mail Station E-175
Hughes Aircraft Company
Culver City, California 90230

Director, USAF PROJECT RAND
Via: Air Force Liaison Office
The RAND Corporation
Attn: Library D
1700 Main Street
Santa Monica, California 90406

Deputy Director and Chief Scientist
Office of Naval Research Branch Office
1030 East Green Street
Pasadena, California 91101

Aeronautics Library
Graduate Aeronautical Laboratories
California Institute of Technology
1201 E. California Blvd.
Pasadena, California 91109

Professor Nicholas George
California Inst. of Technology
Pasadena, California 91109

Commanding Officer
Naval Weapons Center
Corona Laboratories
Attn: Library
Corona, California 91720

Dr F. R. Charvat
Union Carbide Corporation
Materials Systems Div.
Crystal Products Dept.
8888 Balboa Avenue
P.O. Box 23017
San Diego, Calif 92123

Hollander Associates
P. O. Box 2276
Fullerton, California 92633

Commander
U. S. Naval Missile Center (56322)
Point Mugu, California 93041

W. A. Eberspacher, Associate Head
Systems Integration Division
Code 5340A, Box 15
U. S. Naval Missile Center
Point Mugu, California 93041

Sciences-Engineering Library
University of California
Santa Barbara, California 93106

Commander (Code 753)
Naval Weapons Center
Attn: Technical Library
China Lake, California 93555

Library (Code 2124)
Technical Report Section
Naval Postgraduate School
Monterey, California 93940

Glen A. Myers (Code 52Mv)
Assoc. Professor of Elec. Engineering
Naval Postgraduate School
Monterey, California 93940

Dr Leo Young
Stanford Research Institute
Menlo Park, Calif. 94025

Lenkurt Electric Co., Inc.
1105 County Road
San Carlos, California 94070
Attn: Mr E. K. Peterson

Director
Microwave Laboratory
Stanford University
Stanford, California 94305

Director
Stanford Electronics Laboratory
Stanford University
Stanford, California 94305

Director, Electronics Research Laboratory
University of California
Berkeley, California 94720

DELETE:

LTC H. W. Jackson

REPLACE WITH:

Dr L. A. Wood, Director
Electronic & Solid State Sciences
Air Force Office of Scientific Res.
1400 Wilson Boulevard
Arlington, Virginia 22209

DELETE:

Mr Norman J. Field

REPLACE WITH:

Mr Norman J. Field, AMCPM-AA-PM
Chief, Program Management Division
Project AACOMS, USAECOM, Bldg 2525
Fort Monmouth, New Jersey 07703

DELETE:

Mr Robert O. Parker, AMSEL-RD-S

REPLACE WITH:

Mr I. A. Balton, AMSEL-XL-D
Executive Secretarty, TAC/JSEP
U. S. Army Electronics Command
Fort Monmouth, New Jersey 07703

ADDENDUM:

Dr H. K. Ziegler, Chief Scientist
Army Member TAC/JSEP (AMSEL-SC)
U. S. Army Electronics Command
Fort Monmouth, New Jersey 07703

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

University of Illinois
Coordinated Science Laboratory
Urbana, Illinois 61801

2a. REPORT SECURITY CLASSIFICATION

2b. GROUP

3. REPORT TITLE

THE IMPLEMENTATION OF A RELATIONAL DOCUMENT RETRIEVAL SYSTEM

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*

Yean-Hsi CHANG

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September, 1970	48	9

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
DAAB 07-67-C-0199; also in part OE C-1-7-071213-4557.	R-489
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	UILU-ENG-70-234

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Joint Services Electronics Program thru U. S. Army Electronics Command Fort Monmouth, New Jersey 07703

13. ABSTRACT

A relational document retrieval system is developed that is primarily concerned with hierarchic tree structures and lateral links. The lateral linkages are relationships that cannot be represented within the tree structure representation. A number of annotated examples are given along with a detailed description of the relational data structure and the programs that manipulate and search this structure. The system was programmed in ISL, a string manipulating language designed for information retrieval purposes.

DD FORM 1 NOV 65 1473

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Relational Structure Lateral Links Hierarchic Tree Structure Coding Theory Information Searching Language Structure Representation						